ESSAYS ON HOUSING AFFORDABILITY, CHILD DEVELOPMENT AND ECONOMIC MOBILITY

by

Hanchen Jiang

A dissertation submitted to The Johns Hopkins University in conformity with the requirements for the degree of Doctor of Philosophy.

Baltimore, Maryland

June, 2020

Abstract

This dissertation consists three essays on the economics of housing affordability, child development, and economic mobility. In these essays, I explore the unintended policy consequences of rent regulation on tenant labor market outcomes, the intergenerational effect of the Earned Income Tax Credit on child development, and how paternal migration affects left-behind children's non-cognitive development in rural China.

Chapter One focuses on the rent regulation, which is central to the affordable housing policies of local municipalities and is on the rise in the United States and worldwide. In this essay I explore the unintended consequences of rent regulation on tenant labor market outcomes, along with the impact that policy awareness has on those outcomes, using a novel data set on rent stabilization in New York City. Recognizing the potential endogeneity of living in a rent-stabilized unit, I construct an instrumental variable that leverages variation in the availability of rent-stabilized units across New York boroughs over three decades of data. I find that rent-stabilized tenants are more likely to be unemployed compared with tenants in private market-

rate units, particularly among white and high-skilled tenants. Furthermore, I identify policy awareness using a unique feature of the data, and show that a large share of rent-stabilized tenants are either misinformed or unaware of their rent regulation status. The impact of rent stabilization on unemployment only exists among tenants who are aware of their regulation status.

Chapter Two aims to investigate the impact of the Earned Income Tax Credit (EITC) on children's cognitive and non-cognitive skills. Given that the theoretical impact of the EITC on child development is ambiguous, this essay particularly emphasizes the underlying mechanisms. Using data from the Panel Study of Income Dynamics (PSID) and its Child Development Supplement (CDS) in combination with an augmented sample of all non-parental child care histories, increased EITC generosity is found to negatively affect cognitive skills of children from single-mother families. No significant effect on children from married-mother families is found. Positive effects are found regarding children's non-cognitive skills, irrespective of family structure. Furthermore, several mechanisms are shown to be important. First, there is strong evidence that both single-mother and married-mother families increase their non-parental child care usage when facing higher EITC generosity, but this is enacted in different forms. Single mothers mainly utilize relatives for child care, while married mothers tend to utilize formal child care programs. Second, single mothers tend to shift from actively engaged time with children to passively accessible time. Third, single mothers are found to have increased psychological distress. Both the time input

and psychological distress channels are not salient among married-mothers. Finally, the analysis shows that increased goods inputs, measured by income, may not directly correspond with increased child-specific expenditures.

In Chapter Three we look at how paternal migration affects left-behind children's non-cognitive development in developing countries. We use longitudinal data of children in rural China and adopt labor market conditions in destination provinces as instrumental variables for paternal endogenous migration choice. We find that father's migration has a significant negative effect on children's non-cognitive development, particularly for boys. We test and discuss several mechanisms including parental financial inputs, parental time inputs, children's own time input, household bargaining, and parenting style.

Advisors

Professor Robert A. Moffitt (Primary Reader and Adviser)

Professor Yingyao Hu (Secondary Reader)

Acknowledgments

I would like to thank my Ph.D. advisers, Professor Robert Moffitt, Yingyao Hu, Richard Spady, and Luis E. Quintero for their guidance, support, and encouragement throughout my Ph.D. time. I would like to express my deepest appreciation to Professor Moffitt and I feel extremely lucky to have him as my intellectual role model and primary adviser. Professor Moffitt has always encouraged me to work on questions that I am interested in, supported me in terms of applying to data, grant, and summer schools, and offered insightful suggestions and guidance throughout the research journey. I am also particularly grateful to Professors Yingyao Hu, Richard Spady and Luis E. Quintero for their tremendous encouragement and invaluable advice.

I would also like to thank many other professors, mentors, and co-authors who have also kindly offered incredible support and help: Kathryn Anderson (my M.A. adviser at Vanderbilt University), Burt Barnow, David Bishai, Shahe Emran, James Foster, Daniel Gubits, Lingxin Hao, Chris Herbst, Matthew Kahn, Elena Krasnokutskaya, Sandra Newman, Nicholas Papageorge, Edral Tekin, Forhad Shilpi, Xi Yang, among others.

ACKNOWLEDGMENTS

I would also like to thank my friends and colleagues: Sung Ah Bahk, Ruoyu (Steven) Chen, Dong (Carl) Cheng, Edmund Crawley, Luigi Durand, Wenzheng Gao, Yajing Jiang, Osama Khan, Delong Li, Yizhou (Andrew) Liu, Chang Ma, Victor Ronda, Rich Trapp, Shawn Wu, Biancen Xie, Yi Xin, Jiaxiong Yao, Shiqi Wang, Kevin Yuan, Tong Zhou, among others. They made the long Ph.D. process fun and enjoyable.

A special thanks to Dr. Jean Kim from the Department of Otolaryngology at Johns Hopkins Bayview Medical Center, who provided the nasal sinus surgery for me in November 2017 with her extraordinary expertise and kindness.

Finally these essays would not have been possible without the unconditional love and never-ending support of my parents and my wife. I would not have had the courage to start my Ph.D. journey without my parent's encouragement. It was the Johns Hopkins that made me meet my wife. They always show faith in me, especially when I was lacking it myself.

Dedication

This thesis is dedicated to my parents, Xueping and Weiping, and my wife, Danjing.

Contents

CONTENTS

List of Tables

LIST OF TABLES

List of Figures

Chapter 1

Rent Regulation, Policy Awareness, and Labor Market Outcomes: Evidence from New York City

[1]

[1]I would like to express my great appreciation to Robert Moffitt, Yingyao Hu, Matthew Kahn, Luis Quintero, and Richard Spady for their encouragement, guidance, and support. I would also like to also acknowledge invaluable suggestions from Ingrid Gould Ellen and Katherine O'Regan. For other useful comments and thoughtful conversations, I would like to thank the following: Brent Ambrose, Burt Barnow, Jaclene Begley, Paul Carrillo, Edmund Crawley, Edward Coulson, Vadim Elenev, Edward Glaeser, Daniel Gubits, Chris Herbst, Nicholas Hill, Benjamin Keys, Sandra Newman, Hal Martin, Raven Molloy, Gary Painter, Mark Partridge, Nicholas Papageorge, Stuart Rosenthal, Bryan Stuart, Erdal Tekin, Rich Trapp, Stijn Van Nieuwerburgh, Xi Yang, and Yves Zenou, along with participants at the 2018 Workshop on Urban and Regional Economics at Singapore Management University, 2019 American Real Estate and Urban Economics Association (AREUEA) Doctoral Session and National Conference, 2019 Eastern Economics Association Annual Conference, 2019 American Real Estate Society (ARES) Doctoral Session, 2019 Conference on Real Estate and

1.1 Abstract

Rent regulation is central to the affordable housing policies of local municipalities and is on the rise in the United States and worldwide. In this paper I explore the unintended consequences of rent regulation on tenant labor market outcomes, along with the impact that policy awareness has on those outcomes, using a novel data set on rent stabilization in New York City. Recognizing the potential endogeneity of living in a rent-stabilized unit, I construct an instrumental variable that leverages variation in the availability of rent-stabilized units across New York boroughs over three decades of data. I find that rent-stabilized tenants are more likely to be unemployed compared with tenants in private market-rate units, particularly among white and high-skilled tenants. Furthermore, I identify policy awareness using a unique feature of the data, and show that a large share of rent-stabilized tenants are either misinformed or unaware of their rent regulation status. The impact of rent stabilization on unemployment only exists among tenants who are aware of their regulation status.

Housing at Fisher School of Business at Ohio State University, 2019 Workshop on Urban and Regional Economics at Bogotá, Colombia, 2019 National Association for Business Economics (NABE) Tech Econ Conference, Johns Hopkins Carey Business School, and Vanderbilt University. Research funding from the 21st Century Cities Initiative Award for Doctoral Research on Urban Issues at Johns Hopkins University is appreciated. All errors are my own.

1.2 Introduction

Rent regulation is central to the affordable housing policies of local municipalities. Its modern version, which limits rent increases, has recently regained legislative momentum as a result of the rental housing affordability crisis. For example, Oregon, New York, and California all passed rent regulation bills in 2019. Many more states in the United States are also considering bringing back rent regulation, with some presidential candidates even calling for national rent control.[2] In this paper I explore the hitherto understudied effect of rent regulation on labor market outcomes.

In particular I answer two questions: what are the impacts of rent regulation on tenant labor market outcomes, and would the impacts be different depending on policy awareness - whether rent-regulated tenants correctly know their rent regulation status? While there have been extensive studies evaluating the policy's impacts on housing price, supply, and quality from the seminal work by Olsen (1972) to the recent paper by Autor et al. (2014),[3] much less is known about the policy's spillover effects on behavioral outcomes. Interest in spillover effects has grown recently, where the outcomes of interest vary from crime (Autor et al., 2017) to eviction (Asquith, 2017) to tenant mobility, displacement, and inequality (Diamond et al., 2019).[4] However,

[2]Outside the United States, Berlin has introduced a five-year rent freeze plan to counter rising rents in 2019 as well. A rent freeze is usually regarded as first generation rent regulation, commonly known as rent control. Paris also reintroduced rent control earlier in 2019.

[3]Other empirical studies include Early (2000); Early and Olsen (1998); Gyourko and Linneman (1990); Moon and Stotsky (1993); Nagy (1997); Olsen (1988); Sims (2007) and others. For excellent reviews, see Arnott (1995); Metcalf (2018); Turner and Malpezzi (2003) and Pastor et al. (2018).

[4]Most studies are within the partial equilibrium framework in nature. The only exception is Favilukis et al. (2019), who build a general equilibrium model and discuss the role of different

the effect of rent regulation on tenant labor market outcomes in the United States remains an unexamined question. Rent regulation matters for labor market outcomes because it generates a rent subsidy for rent-regulated tenants. Such a rent subsidy not only offers insurance value and eases the intertemporal budget constraint, but it also induces "lock-in" by reducing residential and geographical mobility, both of which could affect labor outcomes, such as job search behavior and unemployment.[5]

One of the main contributions of this paper is to propose and quantify a novel causal claim of rent regulation on tenant labor market outcomes. Researchers have studied the spillover effects of other assisted housing programs on recipients' labor market outcomes, such as Van Dijk (2019); Jacob and Ludwig (2012); Newman et al. (2009); Shroder (2002); Susin (2005) and others.[6] However, rent regulation differs in critical ways from most assisted housing programs that have been studied, in particular, public housing and housing voucher programs. Specifically, rent regulation is not a means-tested welfare program targeted at the low-income population. It has no income eligibility rule, no formal application requirement, and no work requirement, nor does the benefit decrease when a renter's earnings increase. From an ex-ante perspective, rent regulation acts like a universal in-kind transfer in which tenants

affordable housing policies on city welfare.

[5]To the best of my knowledge, the only paper that also links rent regulation with labor market outcomes is Svarer et al. (2005), which looks at rent control in Denmark. However, the policy institutions and the rental housing markets are fundamentally different between the United States and Denmark. Moreover, to reach causation, my paper pays close attention to the self-selection in rent-regulated housing, which is likely to be an endogenous choice.

[6]Olsen (2003) and Collinson et al. (2015) offer excellent reviews on the effect of assisted housing programs on various outcomes.

with all backgrounds are eligible to participate. Hence, not only is the overall effect of rent regulation on tenant labor market outcome important, but so is the evidence on the ex-post incidence of the program and its potential heterogeneous effects. Such evidence sheds light on the equity, efficiency, and potential misallocation of rent regulation policy, which have been emphasized in Gyourko and Linneman (1989), Glaeser (1996) and Glaeser and Luttmer (2003).

This paper also offers empirical evidence about policy awareness and highlights its important role in affecting tenant labor market outcomes. Program participants are commonly assumed to be rational in the program evaluation/treatment effect literature. This assumption suggests that rent-regulated tenants are fully aware of their regulation status and, accordingly, their rent subsidy. However, I document new evidence to verify this assumption and explore its implication. Policy awareness matters because it affects tenants' forward-looking decision-making - tenants who are unaware of their status will not behave in this way.

This paper focuses on rent stabilization in New York City, in which rent increases are limited.[7] The empirical analysis uses the New York City Housing and Vacancy Survey (NYCHVS). This data set is representative for the entire city and is conducted about every three years by the U.S. Census Bureau. It contains detailed information on both housing and household characteristics. A unique feature of this data set is it provides accurate rent regulation status of each housing unit verified by administrative

[7]In 2017, New York City had about 1 million rent-stabilized housing units - even more than private market-rate units.

sources, irrespective of being occupied or vacant.

There are endogeneity concerns in assessing the causal impact of rent regulation on tenant labor market outcomes, because renters living in rent-stabilized units might be systematically different in unobserved ways from those living in private market-rate units. Those differences, such as individual skills, preferences, or labor market shocks, might also directly affect tenant labor market outcomes but be unobservable to researchers. Therefore, any statistically significant association does not necessarily convey causation. I first provide evidence that rent-stabilized tenants are observationally similar to tenants in private market-rate units, particularly in terms of education and income. This pattern is distinct from that is commonly found in other assisted housing program studies in which program beneficiaries often have lower socioeconomic background and educational attainment.

I use an instrumental variable (IV) strategy to overcome the possible self-selection concerns. Equipped with three decades of housing and vacancy data since 1978, I achieve causal inference by leveraging the variation in the availability of rent-stabilized units across New York boroughs over time. To strengthen the validity of the instrument, various procedures are applied to alleviate different concerns such as unobserved confounding factors at the borough level, move-in year endogeneity, and model misspecification. Last but not least, I apply the sorted effect method, recently developed by (Chernozhukov et al., 2018), to discover heterogeneous effects beyond their averages.

Furthermore, a unique feature of the data allows identification of policy awareness because the data contain information about both the legal rent regulation status and tenants' self-perception on all units. This paper also leverages local rent-regulation knowledge differences at the sub-borough level (inspired by Chetty et al. (2013)) as an additional instrument for policy awareness to strengthen the causation and to investigate whether policy awareness matters for the impacts of rent stabilization on tenant labor market outcomes.

Both ordinary least squares (OLS) and probit baseline estimates show that rent-stabilized tenants are more likely to be unemployed (about 2.3-3.5 percentage point) when compared with renters in private market-rate units. The IV estimates confirm that the negative effect on unemployment is indeed causal. Moreover, the estimated coefficients are larger under IV estimation when compared with OLS and probit baseline estimation. This suggests that the self-selected tenants may have better characteristics, which is unobservable and leads to underestimating the negative effect in baseline estimation[8]. Moreover, the effects are heterogenous and possibly counterintuitive: the unemployment effects are only significant among high-skilled, white tenants, while no significant effects having been found among low-skilled, non-white tenants. There are no significant causal impacts regarding other labor market outcomes, such as labor force participation or hourly wage rate.

This paper also documents a new empirical fact about the policy awareness of

[8]Possible examples of such characteristics include strong search skills or local networks that help tenants to find both rent-stabilized units and jobs more easily.

rent stabilization: not everyone knows about it! Among all rent-stabilized tenants, I find that only 34 percent of those who live in rent-stabilized units are able to correctly identify their rent regulation status. In contrast, more than 24 percent are misinformed, believing their housing units are not regulated at all despite living in rent-stabilized units.[9] Descriptive analysis shows that aware renters have a much better socioeconomic background compared with renters in private market-rate units, while misinformed renters have a much worse socioeconomic background compared with renters in private market-rate units.

Results based on the IV strategy demonstrate that rent-stabilized tenants who are aware of their rent regulation status are more likely to be unemployed, when compared to renters in private market-rate units. However, no significant causal impact is found when comparing rent-stabilized tenants who are misinformed with renters in private market-rate units. Such findings may help readers understand why the effect of rent stabilization on unemployment is only seen among high-skilled tenants: the aware group has a much higher education attainment even when compared with private market renters. High-skilled workers are commonly observed to have higher geographical mobility than low-skilled workers.[10] Therefore, rent stabilization may be expected to have a greater effect on the aware, high-skilled tenants who are "locked-in" in a superstar city like New York City (Gyourko et al., 2013). This issue is also

[9] In addition, 22 percent do not know about regulation status and another 20 percent refused to answer.

[10] See Austin et al. (2018); Bound and Holzer (2000); Ganong and Shoag (2017); Notowidigdo (2019); Topel (1986) among others.

broadly related to the Oswald hypothesis, which postulates the negative effect of homeownership on unemployment and other labor market outcomes (Blanchflower and Oswald (2013); Oswald (1999, 1996)).[11]

The rest of the paper proceeds as follows: Section 2.3 discusses the institutional background. Section 2.5 introduces the data and the descriptive analysis. Section 1.5 offers theoretical discussion and then elaboration on the empirical strategy. Section 1.6 reports the main results about the effect of rent regulation on labor market outcomes. Section 1.7 extends the discussion to the role of policy awareness. Section 1.8 offers evidence on suggestive mechanisms and checks the robustness. Section 2.10 discusses policy implications and conclusions.

1.3 Institutional Background

Rent stabilization is the modern version of rent regulation, which began in New York City in 1969. Rent stabilization protects tenants in two major ways: (1) tenants are protected from sharp rent increases; (2) tenants have the right to renew their leases and are protected against arbitrary evictions.[12] It is far more popular today than its predecessor - rent control, which is also known as the first generation of rent regulation (Arnott, 1995). Rent control involves more stringent restrictions on rent

[11] Several studies have found negative effects of homeownership on unemployment and wages across multiple countries and periods (Belot and van Ours, 2001; Coulson and Fisher, 2002; Green and Hendershott, 2001; Munch et al., 2006; Di Tella and MacCulloch, 2005; Yang, 2019).

[12] This provides more protection against eviction than in a private market-rate apartment. However, eviction may occur if a rent-stabilized unit is not used as a primary residence, among other situations.

increases (even freezing rents in extreme cases), but it has dwindled in popularity due to rent-controlled units either become rent-stabilized or completely deregulated upon vacancy. In 2017, of the 2 million rental market units in New York City, there were 966,000 rent-stabilized apartments while only about 22,000 rent-controlled apartments.[13] This paper focuses exclusively on rent stabilization.[14] Figure 1.1 shows the spatial distribution of rent-stabilized units in the rental market.[15]

Eligibility of Rent Stabilization. In New York City, rent stabilization generally applies to three types of apartments in buildings of: (1) six or more units built between February 1, 1947, and January 1, 1974; (2) six or more units built before February 1, 1947, and tenants who moved in after June 30, 1971; and (3) three or more apartments constructed or extensively renovated since 1974 with special tax benefits, such as J-51, 421a, or other programs. The third category only applies to the period when tax abatement is effective, which usually lasts for 10-20 years. The criteria are shown in Figure 1.2, wherein the shadowed areas represent rent-stabilized units.

Rent Determination. The annual rent increases of rent-stabilized units in New York City are capped and adjusted by the Rent Guidelines Board (RGB). For ex-

[13]Rent control in New York City is regarded as a dwindling stock occupied by an elderly, low-income population. The median age of rent-controlled tenants was 70 in 2011, and the median annual total household income in 2013 was less than $30,000 (in 2016 dollars).

[14]A brief comparison between these two types of rent regulation is summarized in Table 1.23. For a complete history of rent regulation in New York City, readers can refer to `https://www1.nyc.gov/assets/rentguidelinesboard/pdf/history/historyoftheboard.pdf`.

[15]The spatial distribution of rental units and private market-rate units in the rental market are shown in Figures 1.10 and 1.11.

ample, between October 1, 2019, and September 30, 2020, a maximum 1.5 percent increase is allowed for a one-year lease. This guideline has been re-set and has fluctuated every year since 1969.[16] Such adjustment is calculated based on the following criteria: consideration of housing market conditions, economic trends, owner costs and revenues, availability of financing, changes in housing supply, affordability of rental housing, rental vacancy rates, cost-of-living indices and other factors. The main impact of rent stabilization is that this annual adjustment is *lower* than the market rent growth. Figure 1.3 documents the average monthly contract rent for private market-rate and rent-stabilized units in New York City. Clearly, rent increases are significantly smaller for rent-stabilized units compared with private market-rate units.

Greater rent increases are allowed in two situations: (1) when certain types of capital improvements are undertaken in the unit or large investments are put into the building and (2) when turnover (known as "vacancy bonus" or "vacancy allowance") occurs if an unit is still subject to rent regulation.

Deregulation. There are several main cases for when a rent-stabilized unit is deregulated. Most commonly, a rent-stabilized unit is deregulated when the rent reaches the deregulation rent threshold (DRT)[17] and when one of the following two conditions are met: (1) there is a vacancy, or (2) the household's income is above the

[16]The historical maximum increase approved by RGB is documented in Table 1.24.

[17]Between June 19, 1997, and June 23, 2011, the DRT was $2,000. In 2019, the DRT was $2,774.76 for New York City. More details can be found at `https://hcr.ny.gov/system/files/documents/2018/10/deregulationrentincomethreshold.pdf`.

deregulation income threshold (DIT).[18] A rent-stabilized unit may also be deregulated upon vacancy during the conversion to a co-op or condo. Lastly, deregulation may occur when tax abatements or exemptions that the owner receives expire (for example, J-51 or 421-a tax benefits).[19]

Comparisons with Federal Assisted Housing Programs. Rent regulation is special because it differs in critical ways from most assisted housing programs that have been studied at the federal level.[20] Specifically, rent regulation is not a means-tested welfare program targeted at the low-income population. As a result, rent-stabilized tenants have a much wider income distribution. In contrast, most assisted housing programs have certain forms of income eligibility rules. For example, public housing tenants can technically earn up to 80 percent of the area median income (AMI), while such numbers have been lower for both the low-income housing tax credit (LIHTC) program (60 percent) and housing choice voucher program (50 percent).[21] Also, most assisted housing programs require participants to pay about 30 percent of their adjusted income, such as the public housing or housing choice voucher programs. In contrast, the rents of rent-stabilized units are not based on tenants' income at all, which is more similar to the LIHTC program with a flat rent scheme. Lastly,

[18]Between July 1, 1998, and June 30, 2011, the DIT income was $175,000. From July 1, 2011 onward, the DIT has remained at $200,000. See `https://hcr.ny.gov/system/files/documents/2018/10/deregulationrentincomethreshold.pdf.` for more information

[19]Rental units in buildings currently receiving J-51 or 421-a benefits are required to be stabilized, regardless of whether the rent level exceeds the high rent deregulation level.

[20]Regarding U.S. federal assisted housing programs and the related policy discussion, see Collinson et al. (2015); Green and Malpezzi (2003); Gyourko and Glaeser (2008); Olsen (2003) for excellent reviews.

[21]See Table 2.1 in Collinson et al. (2015).

unlike other assisted housing programs, rent-stabilized units do not require any formal application process. Many tenants obtain their rent-stabilized units through good fortune. If an apartment is indeed rent-stabilized, the landlord is supposed to attach to the lease the rent stabilization "lease rider." The rider informs rent-stabilized tenants of their rights and responsibilities. However, it is not clear how strictly such requirements are enforced in reality given it is the last step of signing the lease. One of the novel empirical findings of this paper is that many rent-stabilized tenants are indeed not aware of their good fortune.

Rent-Stabilized Units Are Difficult to Find. Rent-stabilized units are difficult to find for the following reasons. First, although a building may contain rent-stabilized units, not all units in the same building have the same rent-stabilization status. Moreover, landlords might not advertise rent-stabilized units publicly. As shown in Figure 1.4, less than 3 percent of online postings on the `StreetEasy` website contain advertisements related to rent stabilization.[22] However, historical NYCHVS data suggests that at least 25 percent of vacant-for-rent units are rent-stabilized, which makes searching for rent-stabilized apartments more difficult and often requires specialized knowledge or local networks.

[22]The `StreetEasy` website (`https://streeteasy.com`) is unique because it allows users to directly search for keywords in addition to select various criteria. An online posting is considered to have rent stabilization advertisement if any of the following words appear in the description section: rent stabilization, rent stabilized, stabilized, and some others. Other websites do not even include rent stabilization information at all.

1.4 Data and Descriptive Analysis

1.4.1 The New York City Housing and Vacancy Survey (NYCHVS)

This paper takes advantage of a novel data set that is underexplored - the New York City Housing and Vacancy Survey (NYCHVS) - for the empirical analysis. It is representative of the entire city and covers all the five boroughs (Bronx, Brooklyn, Manhattan, Queens, and Staten Island). Each wave of data contains approximately 18,000 units, both occupied (either rental-occupied or owner-occupied) and vacant.[23] For occupied units, the survey includes comprehensive information about both housing and household characteristics.

The NYCHVS has a high interview rate of 98 percent and is conducted approximately every three years by the U.S. Census Bureau in compliance with New York state and New York City's rent regulation laws.[24] For this purpose, the rent regulation status of each housing unit is verified through administrative sources, because self-reported receipts are not accurate.[25] This gives NYCHVS a unique angle on answering how rent regulation affects tenant behavior and labor market outcomes.

[23]Each sample unit represents approximately 180 similar housing units. For each occupied housing unit, there is a person-record in addition to a household record for each person in an occupied housing unit. For vacant housing units, a house record is also available.

[24]The survey is sponsored by the New York City Department of Housing Preservation and Development.

[25]As elaborated in a later section, the distinction between legal regulation status and self-reported regulation status plays an important role - it identifies policy awareness.

CHAPTER 1. RENT REGULATION, POLICY AWARENESS, AND LABOR MARKET OUTCOMES: EVIDENCE FROM NEW YORK CITY

This paper uses the pooled 2002 and 2005 waves of NYCHVS to construct the main analytic sample. Almost three decades of data beginning from 1978 are used to construct IVs for causal identification. Previous studies using NYCHVS have almost exclusively focused on the housing market in isolation.[26] One of the contributions of this paper is to take advantage of the rich household information and to answer questions linking housing policy to labor market outcomes.

Moreover, the 2002 and 2005 waves are the most recent surveys that provide information on both actual rent regulation status and tenants' self-perceptions.[27] Such distinction provides a unique opportunity to differentiate between rent-stabilized tenants who are aware of their good fortune and those who are not aware and even misinformed. To the best of my knowledge, this feature has never been used in the literature before.

Sample Restriction. To construct the analytic sample for empirical analysis, I follow these restrictions: (1) Only rental housing units that are either rent-stabilized or private market-rate are kept. This excludes other rental housing units, such as rent-controlled, public housing, and other subsidized rental units, as well as the owner-occupied market. (2) Households benefiting from any federal, state, or city housing

[26]The earliest NYCHVS used in the literature was the 1968 wave (Olsen, 1972), (Gyourko and Linneman, 1989, 1990) and (Ault et al., 1994) in which researchers focused on the role of rent control (instead of rent stabilization) in the housing market. Later waves are used almost exclusively for housing market analysis. For example, the 1978, 1981, 1984, and 1987 data are used in Linneman (1987); Moon and Stotsky (1993); Nagy (1995, 1997) and the 2011 data is used in Sieg and Yoon (2016).

[27]Since 2008, the survey question associated with self-reported perception on rent regulation has been dropped.

subsidy programs are excluded. The main example is the federal Section 8 housing choice voucher program.[28] (3) Households where the household heads' nominal non-labor incomes are greater than $100,000 are excluded.[29] (4) Households where household heads are younger than age 26 or older than 54 are excluded.[30] (5) Households that moved into their current housing units before 1978, due to data limitations, are excluded.[31]

Summary Statistics. Summary statistics on variables related to household and housing are shown in Table 1.1.[32] Detailed variable definitions are provided in **??**.

As shown in panel A of Table 1.1, both rent-stabilized and private market-rate tenants have the same high labor force participation rates (89 percent). These rates are higher than the national average.[33] Rent-stabilized tenants have a relatively higher unemployment rate (6.3 percent) than private market-rate tenants (5.2 percent). The difference here is significant at the 5 percent level.[34] Rent-stabilized tenants are also found to have lower wage rates and annual wage incomes. The differences here are significant at the 1 percent level.

[28] I am grateful for Ingrid Gould Ellen for offering this suggestion.

[29] In 2002 and 2005, there are only 160 and 180 such cases are excluded out of the entire 15,894 and 15,547 rental units respectively.

[30] The empirical results are robust if the upper bound of the age range is extended to 64.

[31] The earliest NYCHVS data that are available are from 1978. IVs are constructed from the time of move-in.

[32] More detailed summary statistics containing standard deviation, minimum, and maximum are available in Table 1.13 and Table 1.14 respectively.

[33] The labor force participation rate among 25- to 54-year-olds was 83.3 percent in 2002, according to the Bureau of Labor Statistics (BLS) (`https://www.bls.gov/opub/ted/2014/ted_20140106.htm`).

[34] National unemployment in 2002 and 2005 was 5.8 percent and 5.1percent, respectively, according to the BLS (`https://data.bls.gov/timeseries/LNU04000000?periods=Annual+Data&periods_option=specific_periods&years_option=all_years`).

Regarding individual characteristics, educational attainment is fairly balanced between rent-stabilized and private market-rate tenants. In terms of demographics, household heads living in rent-stabilized units are more likely to be female and slightly older. They are less likely to be born in New York City, to be married, and to have children, and they are more likely to have smaller families. In terms of ethnicity, household heads living in rent-stabilized units are less likely to be white or Asian (but more likely to be Hispanic). There is no significant difference regarding African-Americans. Lastly, there is no observational difference in terms of the household heads' non-labor incomes. In the next section, a regression analysis is conducted to further investigate what individual characteristics are predictive of being a rent-stabilized tenant or not. Geographic differences at the sub-borough level, which are not revealed here in the summary statistics comparison, are particularly considered.

Housing characteristics are shown in panel B of Table 1.1. First, the monthly contract rent and monthly gross rent of rent-stabilized units are significantly smaller than those of private market-rate units. The average rent difference is approximately $300 per month. Second, the tenure of residence is much longer (more than 2 years) for tenants living in rent-stabilized units than for those in private market-rate units. Furthermore, rent-stabilized units are generally smaller than private market-rate units measured by number of rooms and bedrooms. Last, rent-stabilized units seem to be of relatively lower quality at both the unit and building level. In the next section, a regression analysis is conducted that takes such quality difference into account.

1.4.2 Is Rent Stabilization Valuable?

One may ask if rent stabilization is valuable. When the regulated price is below the market equilibrium price, landlords can financially benefit from reducing housing quality in rent-regulated properties. In a static world, assuming housing quality is linear in price, a landlord may be fully compensated for the wedge between equilibrium price and regulated price by reducing housing quality sufficiently. In such a scenario, tenants living in rent-stabilized units do not benefit at all.

Evidence from Rent Stabilization Dummy. Following Autor et al. (2014), I estimate

$$P_{ijst} = \alpha + \beta \text{ RS}_i + \lambda \text{ RS}_i \times \text{Tenure}_i + \gamma X_{ijbt} + \epsilon_i \tag{1.1}$$

where P_{ijbt} indexes the monthly rental price (either in level or in log) of a housing unit i in building/neighborhood j located in sub-borough s at year t. RS_i indicates whether a housing unit is rent-stabilized. $\text{RS}_i \times \text{Tenure}_i$ captures how the value of rent stabilization is larger if one stays in the same dwelling longer. X_{ijst} is a vector that contains different levels of housing and building traits.

I first look at the effect of rent stabilization at the level of monthly contract rents. The result is documented in panel A of Table 1.2, wherein different levels of housing traits are progressively added. Overall, rent-stabilized units do have lower quality-adjusted rents. The rent discount also increases with a longer tenure of the residence. On average, the monthly quality-adjusted contract rent is more than $350 per month for a newly moved-in tenant. Every additional year leads to an approximately $ 20

additional discount per month. For someone who has lived in the same unit for 10 years, the average rent discount could be more than $550 per month.

The log of monthly contract rents is also considered as an alternative outcome variable. The results are shown in panel B of Table 1.2. The general takeaway is the same as in the previous paragraph: rent stabilization is valuable after adjusting for quality differences. Without having duration of tenure, rent stabilization leads to a 24 percentage point reduction in contract rent. After adding duration of tenure, rent stabilization leads to a 15 percentage point reduction in contract rent. In addition, one additional year is linked with a further 1.8 percentage point decrease in rent.[35]

Evidence from Hedonic Rent Discount. An alternative approach to investigating the value of rent stabilization is to estimate the rent discount directly, which allows for greater heterogeneity beyond average. The rent discount is defined as the price difference between the actual contract rent and the predicted counterfactual market rent for a rent-stabilized unit. The predicted counterfactual market rent is calculated using a hedonic pricing approach in the spirit of Rosen (1974).[36] Detailed discussion on how this rent discount is calculated can be found in Jiang et al. (2019). From this related study, I can draw the estimated rent discount directly. In Figure 1.12, the rent discount (vertical axis) is plotted against the contract rent for rent-stabilized

[35]One might wonder whether there is a nonlinear relationship between rent stabilization and tenure of residence. I have worked on a quadratic interaction term of rent stabilization and tenure of residence, but the result is not significantly different from the current linear interaction specification.

[36]Briefly, this exercise involves a two-step procedure. In the first step, the implicit prices of various housing traits are recovered only by private market units. In the second step, the implicit prices are used to predict the market price for rent-stabilized units.

units (horizontal axis), and a local polynomial smooth line is fitted. It is clear that there exists positive rent discount on average. Furthermore, the rent discount is much larger for stabilized units with lower contract rents. The rent discount also, on average, converges to zero when a rent-stabilized contract rent crosses more than $2,500, which is consistent with the high-rent deregulation policy institution. In addition, Figure 1.13 shows the rent discount against total household income. A flat-fitted line suggests that the overall rent discount is not correlated with family income, which is consistent with the fact that rent stabilization is not means-tested.

1.4.3 Who Lives in Rent Stabilized Units?

Evidence from Observed Characteristics. I first investigate whether rent-stabilized tenants have different socioeconomic and demographic characteristics compared with tenants living in private market-rate units. The answers assist in thinking about the possibly self-selection concern, which is a key question of any empirical policy evaluation.

A regression analysis is conducted where the outcome variable is binary, equaling 1 for rent-stabilized tenants and 0 for private market tenants. The analysis includes a rich set of individual characteristics and uses different model specifications, including the linear probability model (OLS) as well as binary choice models, such as probit and logit models. The results are represented in Table 1.3.[37]

[37]Column A of Table 1.3 is the the linear probability model, which means that coefficients can be interpreted as marginal effects directly. Columns B and C of Table 1.3 are probit and logit models,

Most individual characteristics are *not* significantly associated with living in rent-stabilized units, particularly with respect to education and income (either the household head's own total income or the total income of all the other family members). This pattern is distinct from other assisted housing programs, such as public housing or housing voucher programs in which tenants with lower education and income are most likely to participate. Many other characteristics prove insignificant, such as age, coresidence with parents, birth place, and others. The only distinct factors are gender and family size.[38] These findings are robust to alternative binary choice models, such as probit and logit.

Sources of Endogeneity. Even though tenants in rent-stabilized units seem to be observationally similar to tenants in private market-rate units, one may still be concerned with other sources of endogeneity that are unobservable to researchers but affect tenant labor market outcomes directly, that is, beyond variables included in Table 1.3.

Let's first revisit the canonical skill (or ability) bias, which is arguably the most classical source of omitted variable bias in labor economics.[39] On the one hand, given that rent-stabilized units have lower rents on average, it is plausible that tenants with

respectively. The coefficients have been transformed as marginal effects to ease the interpretation and comparison to a linear probability model.

[38]In addition to the baseline model, I also further explore whether there is any pattern associated with a specific industry or occupation category. While no significant pattern associated with industry types has been found, there is some level of heterogeneity associated with occupation. For example, being a social scientist or lawyer is negatively associated with living in rent-stabilized units, while being a physical scientist or cleaning and building service worker is positively associated with living in rent-stabilized units. Results are available upon request.

[39]For example, the wage return to schooling literature summarized in Card (2001), among others.

lower unobserved skills may have stronger incentives to search for cheaper housing units. Such tenants with lower unobserved skills may also have worse labor market outcomes, irrespective of living in rent-stabilized units or not. In this way, a naive OLS estimation may be overestimating, if rent stabilization is found to be negatively associated with labor market outcomes. On the other hand, given that rent-stabilized units are difficult to find, it is plausible that tenants with higher unobserved skills may more easily locate such units. For example, such tenants may be good at online searching or have a strong local network, which may also help with job search and improve other labor market outcomes. In this way, a naive OLS estimation may be underestimating the true effect, and thus, theoretically ambiguous in terms of the direction of bias.

Another source of endogeneity comes from reversed causation. For example, some tenants may expect to have worse labor market outcomes in the future, which causes them to spend more time and effort searching for rent-stabilized units. Hence, an observed negative association between living in rent-stabilized units and labor market outcomes cannot be interpreted as causation either.

A third type of endogeneity stems from unobserved heterogeneity relating to individual preferences and tastes. For example, people who do not want to move and prefer residential stability may spend more time and effort searching for rent-stabilized units. Similarly, people who are more risk averse may also prefer to live in rent-stabilized units so that they do not need to worry about unexpected rent volatility.

In contrast, private market renters may be either more risk inclined or have less preference for stability. The preferences could have effects in the labor market performance as well. Therefore, a significant coefficient associated with rent stabilization may not be causal, and living in a rent-stabilized unit may be a proxy for other preferences and tastes.

In sum, people who expect to gain the most from rent stabilization may place greater effort into searching for a rent-stabilized unit. The factors that drive searching may also be correlated with tenants' future labor market outcomes, which raise concerns related to endogeneity. The next section will address this issue.

1.5 The Effect of Rent Regulation on Labor Market Outcomes

1.5.1 Theoretical Discussion

I first provide theoretical discussion on the effect of rent regulation on labor market outcomes, particularly unemployment and wage.[40] On the one hand, the canonical theoretical framework to understand job search behavior and unemployment goes back to McCall (1970) and Lippman and McCall (1976), where the reservation wage

[40]In the empirical analysis, I also evaluate the effect of rent regulation on labor force participation, given that rent regulation lowers rent and eases intertemporal budget constraint, which may have an income effect.

is endogenously determined and affects unemployment duration in a stationary environment. When rent stabilization is incorporated in such canonical McCall-type job search model, there is no impact of rent stabilization on job search behavior or unemployment duration. The intuition is that for an unemployed worker, no matter whether the worker accepts or rejects a job offer, s(he) would have to pay the exactly the same amount of present discounted value of rents, which is a key property coming out of the stationarity environment.

On the other hand, there are other theoretical alternatives, which are under different assumptions compared with the McCall-type job search model, leading to different predictions.

First, liquidity constraint is not considered in the canonical McCall-type model, which is important in the job search literature, such as (Basten et al., 2014; Card et al., 2007; Chetty, 2008) among others. If unemployed renters are liquidity-constrained, then they would have to accept a job offer earlier than otherwise would be optimal. Rent stabilization may also help ease liquidity constraint by allowing an unemployed renter to search longer for a better job. This would predict that rent stabilization leads to higher unemployment duration.[41]

Second, instead of applying the reservation wage framework, one could also apply the search effort framework, such as (Chetty, 2008; DellaVigna and Paserman, 2005; Paserman, 2008) and others. In this context, the income effects would predict less

[41] Wage is often assumed to be constant in classical job search model with liquidity constraint.

job search intensity, which would also affect unemployment duration accordingly.

Last, spatial job search has not been considered yet, which is more commonly applied in the homeownership and unemployment literature (e.g. Morescalchi (2016); Munch et al. (2006, 2008)) In particular, as illustrated in Coulson and Fisher (2009), different models would have completly different theoretical predictions on unemployment and wage, depending on the specific model assumption. In the end, whether and how rent stabilization affects labor market outcomes demands empirical investigation.

1.5.2 Preliminary Analysis

I start the preliminary analysis by the following baseline empirical set-up:

$$Y_{ijst} = \alpha + \beta \ \mathrm{RS}_{ijst} + \eta X_{it} + \psi V_{jt} + \lambda Z_s + \pi W_t + \epsilon_{ijst} \tag{1.2}$$

where i indexes householder, j indexes the housing unit in a specific building or neighborhood, s indexes the sub-borough, and t indexes years.

Y_{ijst} is the outcome variable of interest. RS_{ijst} is the key variable of interest, which equals 1 when a householder i lives in a rent-stabilized unit j with certain housing traits in sub-borough s and year t.[42] The estimate of interest is β. X_{it} stands for individual-level and household-level characteristics. V_{jt} stands for housing traits at the unit, building, and neighborhood levels. Z_s is the sub-borough fixed effect and W_t is the year fixed effect.

[42] RS_{ijst} equals 0 when a householder lives in a private market-rate unit instead.

The underlying identifying assumption here is that, conditional on regressors X_{it}, V_{jt}, Z_s and W_t, rent stabilization is randomly assigned. As such, it is particularly important to control variables that affect the rent stabilization eligibility. Thus, I include variables such as how many units a building has and when a building is built, that largely affect rent stabilization status.

Different sets of control variables are added progressively. The individual-level demographic and socioeconomic characteristics include gender, ethnicity, marital status, birth place, educational attainment, and nonlabor income (in log). Other household-level characteristics include whether the householder is coresiding with parent(s), whether younger children ($<$ 6 years old) or older children (between 6 and 18 years old) are present, whether non-relatives are present, total family size, householder's parental birth places, and total income of other family members (in log). If the householder is married, then the spouse's educational attainment and labor market status are also controlled. Also controlled are housing traits at different levels, such as the housing unit, building, and neighborhood. The richness of the NYCHVS data also allows me to control for industry and occupation categories, which control those unobserved sector-specific shocks.[43] Last, the sub-borough and year fixed effects are included in all model specifications to allow for unobservable local shocks. Households that have lived in their current units for less than one year are excluded to alleviate reverse causation concern.

[43]Both industry and occupation categories are available not only for those who are currently working but also for those who are currently unemployed or out of the labor force. This excludes, however, people who have been unemployed or out of the labor force for longer than 5 years.

1.5.3 Instrumental Variable Strategy

1.5.3.1 Local Availability of Rent Stabilization as Instrumental Variable

To overcome the endogeneity concern from selection-on-unobservables, I propose an instrumental variable (IV) approach to achieve causation. The goal is to find an instrument Z that affects a tenant's self-selection into a rent-stabilized unit RS but does not affect the error term in the structural equation (Equation 1.2). This study exploits the idea that different boroughs in New York City have different levels of rent-stabilized units that are both vacant and available for rent (vacant-for-rent, hereafter) in different years. This is referred to as the local availability of rent stabilization.

There are two related instruments that capture such variation. The first instrument $\text{IV}^{stab}_{b,t-1}$ is defined as the ratio of the total number of vacant-for-rent units that are rent-stabilized $N^{stab}_{b,t-1}$ and the total number of vacant-for-rent units $N^{all}_{b,t-1}$:

$$\text{IV}^{stab}_{b,t-1} = \frac{N^{stab}_{b,t-1}}{N^{all}_{b,t-1}} \tag{1.3}$$

where b indexes different boroughs (Bronx, Brooklyn, Manhattan, Queens, and Staten Island) and $t-1$ indexes the year before a tenant moves into the current dwelling.

Similarly, I construct another instrument $\text{IV}^{mkt}_{b,t-1}$ as the ratio of the total number of vacant-for-rent units that are private market-rate $N^{mkt}_{b,t-1}$ and the total number of

vacant-for-rent units $N^{all}_{b,t-1}$:

$$\text{IV}^{mkt}_{b,t-1} = \frac{N^{mkt}_{b,t-1}}{N^{all}_{b,t-1}} \tag{1.4}$$

First, it is worth mentioning that the sum of $\text{IV}^{stab}_{b,t-1}$ and $\text{IV}^{mkt}_{b,t-1}$ does not equal to 1 and varies in different years because of the existence of other types of housing, such as rent-controlled housing, public, and other subsidized housing. The total stock of such other types of housing also varies by boroughs and years. $\text{IV}^{stab}_{b,t-1}$ and $\text{IV}^{mkt}_{b,t-1}$ also vary by borough and years.

Second, both the numerator and denominator in Equation 1.3 and Equation 1.4 specifically refer to units that are vacant-for-rent. A vacant unit could either be for selling only or neither for sale nor rent.[44] The differences between vacant-for-rent and vacant have important implications for the validity of the instruments, which are further elaborated on in the next section when discussing the validity of the instruments.

To construct the proposed instruments, one needs to know the following information: (1) when a tenant moves into the current dwelling, (2) how many vacant units there are in each borough at the time 1 year before move-in, and (3) the rent regulation status of all the vacant units in each borough at the time 1 year before move-in.

The data are unique in satisfying all the requirements because NYCHVS depicts a

[44] A unit that is neither for sale nor rent could have the following reasons: (a) is under or awaiting renovation; (b) is to be converted to other types of housing; (c) has a legal dispute; and (d) is being hold for occasional, seasonal, or recreational use.

complete picture of the New York City housing market for more than three decades. Condition (1) is easily fulfilled because such information comes from the main analytic sample in the 2002 and 2005 waves.

Conditions (2) and (3) are more challenging because they require historical data *before* move-in years. The earliest data available from NYCHVS are from 1978, and later waves are as follows: 1981, 1984, 1987, 1991, 1996, 1999, 2002, and 2005. No earlier data currently exist accordingly to the best of my knowledge.[45] In each wave of the NYCHVS, even as early as 1978, there is precise, high-quality information regarding condition (2). This precision exists because, as the U.S. Census Bureau states, the primary policy tool out of NYCHVS is the *vacant available for rent rate*, which is defined as the ratio of the vacant available-for-rent units to the total number of renter occupied and vacant available-for-rent units for the entire city.[46] Regarding condition (3), all rent regulation statuses are also verified through administrative sources, which removes any possibility of self-reported measurement error.

Since the sample of each wave of the NYCHVS is based on the decennial census, which is different for every 10 years, the sampling weights are used to construct $N^{stab}_{b,t-1}, N^{mkt}_{b,t-1}, N^{all}_{b,t-1}$ at the population level to reduce any measurement error coming from different survey methods or sampling designs.[47] Lastly, for any specific year

[45] I are grateful for the confirmation from Joseph Gyourko, Choon-Geol Moon, Edgar Olsen, Janet Stotsky, and the Furman Center at New York University.

[46] Such information is of high quality because the "The design requires the standard error of the estimate of the vacant available for rent rate for the entire city be no more than one-fourth of 1 percent, if the actual rate was 3 percent." See `https://www2.census.gov/programs-surveys/nychvs/technical-documentation/source-and-accuracy/source-2005.pdf` for more information.

[47] For data before 1987, there are no sub-borough identifiers. Hence, the best I can do is construct

that is not available from NYCHVS, I conduct within sampling interpolation using a cubic spline interpolation method.

The variations of proposed instruments $\text{IV}^{stab}_{b,t-1}$ and $\text{IV}^{mkt}_{b,t-1}$ are shown in Figures 1.5 and 1.6 respectively. In both figures, blue dots are directly obtained and calculated from NYCHVS, while the red lines represent fitted value based on cubic spline interpolation. There are clear variations across different boroughs. Moreover, the value of the instruments also varies within the same borough because it is determined by the year of move-in. There is clearly nonlinear variation over time within the borough. For example, in Manhattan, almost 80 percent of vacant-for-rent units were rent-stabilized in 1985. Such percentage decreases over time, becoming less than 40 percent in the late 1990s. However, it started to increase since 1999. Such nonlinear variation shifts the probability of self-selecting into rent-stabilized units and helps to achieve identification.

1.5.3.2 Instrumental Variable Conditions and Discussion of Potential Violations

This section assesses the validity of the proposed instruments so that estimation results can be interpreted as local average treatment effects (LATE) (Imbens and Angrist, 1994). The LATE captures the average causal effect on the subgroup of compliers - tenants whose rent stabilization status is actually affected by the variation

measures at the borough level.

of IVs. The discussion is largely inspired by Bhuller et al. (2018); Caliendo et al. (2017); and Arni and Schiprowski (2019).

Relevance. The relevance condition requires that the instrument significantly affects a tenant's probability of sorting into a rent-stabilized unit given unobservable heterogeneity. This condition is likely to be satisfied and can be easily verified based on the first-stage results. The conceptual discussion uses $\text{IV}^{stab}_{b,t-1}$ for illustration purposes. First, when $\text{IV}^{stab}_{b,t-1}$ is high, then there are relatively abundant rent-stabilized units that are vacant and available-for-rent in a borough's rental market. In that case, whether a household is actively searching for a rent-stabilized unit or not, there is still a relatively higher chance for any household to locate in a rent-stabilized unit. Searching for rent-stabilized housing then becomes much easier and more feasible. The effects change, when $\text{IV}^{stab}_{b,t-1}$ is low. Consider an extreme case in which $\text{IV}^{stab}_{b,t-1} \approx 0$. No matter how strong the self-selection motive is, the chance of successfully locating and renting in a rent-stabilized unit is significantly low. The same logic holds for $\text{IV}^{mkt}_{b,t-1}$, but it affects the probability of a tenant locating in a rent-stabilized unit in the opposite direction.[48]

Exogeneity. The second condition for a valid instrument is exogeneity, meaning that the instrument should not correlate with the error term in the structural equation regarding labor market outcomes (conditional independence). Another in-

[48]That is, a larger value of $\text{IV}^{mkt}_{b,t-1}$ indicates more unregulated market-rate units in a specific borough. This further reduces the likelihood for a tenant to locate and secure a rent-stabilized unit irrespective of the search effort.

terpretation is that the only way through which the proposed instrument affects labor market outcomes is directly through its effect on the endogenous variable (exclusion restriction). For this paper, this condition means that the only way through which local availability of rent stabilization affects tenants' labor market outcomes in the future is by affecting tenants' probability of locating a rent-stabilized unit or not.

First, one must recognize that the proposed instruments are constructed at the borough-level, the variation of which are not affected by individual tenant's housing choice decision one year later. Hence, the instruments are not correlated with sources of endogeneity at the individual level, such as unobserved individual skills, preferences, or labor market shocks.

Second, one may be concerned that the move-in year may be endogenous. This endogeneity is unlikely, and there is direct evidence against such concern. In the data, a tenant moves into a unit for certain reasons, which are reported. Following Newman and Wyly (2006), I classify more than 30 different kinds of reasons into several categories. In particular, I am interested in the employment-related category.[49] There is no significant difference in terms of share of employment-related reasons among all the reasons between rent-stabilized renters and private market renters.

Third, one may also be concerned that the variation of the instruments may be correlated with other borough-level confounders. Such confounders may also persistently affect the labor market in a way that is unobservable to a researcher. On the

[49]The employment-related category contains the following reasons: (a) job transfer or new job, (b) retirement, (c) looking for work, (d) commuting reasons, and (e) other financial or employment reasons.

one hand, I argue that a lot of factors driving the variations of the instruments are independent of the local labor market. The $\text{IV}^{stab}_{b,t-1} = \frac{N^{stab}_{b,t-1}}{N^{all}_{b,t-1}}$ is used for illustration purposes. The denominator, which is total number of *vacant-for-rent* units ($N^{all}_{b,t-1}$), could change even if the total number of *vacant* units remains unchanged. For example, a vacant unit could become available for rent if renovation is completed, a legal dispute is resolved, or is no longer held for seasonal or recreational use. A vacant unit could also become unavailable for rent if the landlord wants to convert it to other types of housing. The numerator, which is total number of *vacant-for-rent* units that are *rent-stabilized* $N^{stab}_{b,t-1}$, could change if (1) a rent-controlled renter moves out or passes away so that the rent-controlled unit becomes decontrolled and rent-stabilized, (2) a renovated or newly constructed apartment receives tax benefits and becomes rent-stabilized, and (3) a previously rent-stabilized unit becomes deregulated for a variety of reasons.

On the other hand, following Sieg et al. (2002), I measure the average unobserved neighborhood quality directly and denote it as $q_{b,t}$, wherein b indexes borough and t indexes time. $q_{b,t}$ is constructed from the following rental pricing function:

$$\ln P^{mkt}_{ibt} = \beta X_{ibt} + \sum_{b=1}^{5} \ln q_{bt} B_{bt} + \epsilon_t \tag{1.5}$$

where P^{mkt}_{ibt} is the rental price (measured as monthly contract rent) for private market-rate unit i in borough b at year t. X_{ibt} contains all the possible *physical housing traits* at the unit, building, and neighborhood level - it does not represent the borough

level. B_t is the borough dummy. Notice that the constant term is not allowed in Equation 1.5, which makes it possible to estimate the coefficient of *all* borough dummies. Heuristically, q_{bt} captures all borough-level unobserved factors, such as school-quality, crime rate, and pollution, some of which may affect future labor market outcomes. These factors are capitalized into the market rental prices that cannot be explained by observable physical housing traits. Most important, the variation of q_{bt} is informative of the changes in unobserved neighborhood quality dynamics on average. q_{bt} is then included as an additional control variable in the IV estimation directly.

Such a procedure is feasible by leveraging the entire 30-year NYCHVS information, which is similar to the way the previous instruments were constructed. In every wave of NYCHVS, I use data at the occupied unit module and estimate Equation 1.5 by focusing only on private market-rate units. For years that are not available in NYCHVS, I also use cubic spline interpolation. The results are shown in Figure 1.7, with plots for both the estimated unobservable borough quality and average monthly rent (in 2017 real value). The points represent actual data observation, and the lines represent cubic spline interpolation. The rental price pattern is consistent with common expectations and further strengthens the validity of the interpolation. Overall, Manhattan has undergone much larger unobserved quality changes when compared with other boroughs.

Monotonicity. The monotonicity condition means that the participation probabil-

ity (in this case, the probability of sorting into a rent-stabilized unit) is a (positive) monotonic function of the instrument, excluding the presence of defiers. It is highly unlikely that tenants would refuse to choose to live in rent-stabilized units when there are larger availability vacant-for-rent units locally, because rent-stabilized units cover a wide range of prices and quality, even including somewhat luxury apartments. Unlike public housing, there is no stigma effect associated with rent-stabilized units. As a result, the monotonicity assumption should be successfully satisfied.

1.5.4 The Sorted Effects Method

Last but not least, this paper applies the sorted effect method developed by Chernozhukov et al. (2018) to understand heterogenous effects beyond their averages. The basic idea of this method is straightforward: the marginal effect in a nonlinear model (such as the probit model) is evaluated for each observation, which is then collected and sorted in increasing order and indexed by percentiles. Instead of only relying on the average marginal effect, this approach completely represents the range of the heterogeneous effects. One of the major technical advances Chernozhukov et al. (2018) make is to offer a bootstrap procedure to calculate the confidence set. This sorted effect method is applied to both an ordinary probit model (in the preliminary analysis section) as well as one augmented with the IV strategy.

1.6 Estimation Results

How does rent stabilization affect tenants' labor market outcomes? First, I report the average effect of rent regulation on different measures of labor market outcomes, based on both OLS estimates and IV estimates, and then I discuss the heterogenous effects.

1.6.1 Baseline Results

OLS Estimates. Before reporting on the causal effect of rent stabilization, I first discuss OLS results as a baseline. The identifying assumption here comes from selection-on-observables. Three sets of outcome variables - labor force participation, unemployment, and wages - are considered. Table 1.4 contains results from the OLS estimations, regressing the outcome variables on rent stabilization dummy while controlling for individual covariates progressively and the sub-borough and year fixed effect.

In panel A, there is no strong association between rent stabilization and labor force participation. The estimated coefficient is not only statistically insignificant but also economically negligible. Moreover, when progressively adding different sets of control variables, the model's explanatory power increases (R^2 becomes larger). This suggests that the insignificance is not because of a lack of overall prediction power in the empirical model. Alternative nonlinear model specifications, such as

probit and logit models, suggest that the result is robust across different models (panel A of Table 1.15), meaning that the insignificant results are not because of the special model assumption.

In contrast, a significant relationship emerges when looking at unemployment status (panel B). Specifically, in column (1), when only the sub-borough and year fixed effects are controlled, living in a rent-stabilized apartment is associated with a 1.5 percentage point higher likelihood of being unemployed. This is statistically significant at a the 1 percent level. However, this model alone is not very informative given the lack of individual covariates. By adding individual demographic, education, and other household-level covariates, the parameter of interest becomes robust in both magnitude and statistical significance. At the same time, the model explanatory power increases. When housing traits in particular are introduced, the coefficient increases to 2.5 percentage point with a significant increase in model prediction power. Last, when all observable characteristics in column (7) are included, the coefficient is 2.3 percentage point and remains significant at the 1 percent level. Based on the insights from Oster (2019), the estimate is robust, and the model prediction power also increases when more individual covariates are introduced. This strengthens confidence in the results that selection-on-observables is also informative for selection-on-unobservables. Alternative model specifications, such as probit and logit models, further confirm that such a relationship is not artificially driven by the linear model specification assumption. Rather, non-linear models such as probit shows a relatively

larger effect: 3.6 percentage point in the probit model compared to 2.3 percentage point in the linear probability model (Ppanel B of Table 1.15).

In panel C of Table 1.4, a similar negative relationship has been found between rent stabilization and hourly wage rate. The coefficient is approximately 7.1. Given the log specification of wage rate, the result is interpreted as rent stabilization is associated with a 7.1 percent lower wage rate.

IV Estimates. I now turn to causal analysis using local availability of rent stabilization as an instrument for a tenant's sorting into a rent-stabilized unit.

First, the first-stage results are in Table 1.5. The main instrument is $\text{IV}^{stab}_{b,t-1}$ (IV1). In column (1) of Table 1.5, the estimation is precise and the sign is expected. The larger $\text{IV}^{stab}_{b,t-1}$ leads to a higher probability of living in a rent-stabilized unit. Specifically, a 10 percent increase of vacant-for-rent and rent-stabilized units in a borough's rental market leads to about a 4.7 percentage point increase of likelihood for living in a rent stabilized unit. The F-statistics in each panel is above 30, irrespective of various outcome variables of interest. The values of this statistic are much larger than the suggested rule-of-thumb threshold of 10 (Staiger and Stock, 1997; Stock et al., 2002a), which suggests a strong first-stage result. In column (2), the instrument is $\text{IV}^{mkt}_{b,t-1}$ (IV2). Similarly, the estimate is significant and the sign is as expected, which means that when there are more private market-rate units among all vacant-for-rent units, a tenant is less likely for a tenant to end up in a rent-stabilized unit. The F-statistic is relatively smaller (around 20) but is still much larger than the

threshold of 10. Last, these two instruments $IV^{stab}_{b,t-1}$ and $IV^{mkt}_{b,t-1}$ are used together (IV3). Given that there is only one endogenous variable and are two instruments, the over-identifying test can be conducted. As shown in column (3), the over-identifying test cannot be rejected in all the panels, where the p-values are between 0.36 and 0.91 in three panels, which further strengthens confidence in the instruments.

I then turn to the main results of rent stabilization on labor force participation, unemployment, and wage rate, which are shown in panels A, B and C of Table 1.6, respectively. Overall, there are significant effects on unemployment (panel B) but no significant results emerge regarding labor force participation (panel A) and wage rate (panel C). The insignificant results with respect to labor force participation is not surprising because the results under IV estimates are actually very similar to the ones under the OLS estimates. The results regarding hourly wage rate are somewhat surprising given that there is a significant negative relationship in the OLS estimates. The estimated sign is still negative, but the standard errors become much larger. This is probably because the hourly wage rate is noisier in such cross-sectional data structure.

I now focus the discussion on unemployment (panel B), wherein both OLS and IV results suggest the same direction and are statistically significant. Panel B of Table 1.6 confirms that rent stabilization indeed causes higher likelihood of being unemployed. The estimates are robust in all three columns and precisely estimated at the 1 percent level. Notably, the IV estimates are larger than the OLS estimates, though both are

statistically significant at the 1 percent level. Under our main instrument (IV1), the point estimate is 0.12. Moreover, the predicted average unemployment rate for rent-stabilized tenants is 8.33% while the predicted average unemployment rate for tenants in the private market-rate units is 2.35%. This suggests an estimated impact of 5.98%.

Why would IV estimates be larger than OLS estimates? This may be possible because of two different types of reasons. First, the self-selected rent-stabilized tenants could have better unobserved characteristics, which makes them less likely to be unemployed in the first place. As mentioned earlier, one possible reason might be that it is not easy to search for rent-stabilized units - it requires strong searching skills that may be correlated with other high (unobserved) skills (see **??** for real suggestions on how to find a rent-stabilized apartment using strong searching skills). Hence, the IV estimates become larger when removing such unobserved skill premiums underlying sorting behavior.

Similarly, Abel et al. (2019) emphasizes the importance of an intention-behavior gap in job search behavior, and the gap is defined as a disconnection between the intention to perform a particular behavior and the enactment of such behavior. As a result, the self-selected tenants may have a less intention-behavior gap (regarding the housing search), which would also be beneficial to their job search in the labor market.

Another possible reason is related to the local network. Anecdotal stories suggest

that the local network is important if a prospective tenant wants to find a rent-stabilized unit. It is well-established that network plays an important role in the job search as well (Cingano and Rosolia, 2012).

The fact that IV estimates are larger than the OLS or probit estimates is also in line with the second explanation, which is related to the interpretation of IV estimates. IV identifies the local average treatment effect (LATE) on the subgroup compliers, meaning those tenants induced to participate in the treatment because of variation in the instruments (Angrist et al., 1996; Heckman, 1997; Heckman and Vytlacil, 2005; Imbens, 2014). This means that the IV identifies the effects among tenants who choose to sort into rent-stabilized units because of larger availability of vacant-for-rent and rent-stabilized units. This may, though, be different from the average treatment effect on all treated tenants (ATT). Nevertheless, the LATE in this case is actually the policy relevant parameter (Heckman and Vytlacil, 2001; Mogstad and Torgovitsky, 2018; Mogstad et al., 2018) because the proposed instruments can be easily affected by policy changes directly.

1.6.2 Heterogenous Treatment Effect

The IV estimates provide the local average treatment effect. How does the causal treatment effect vary among different sub-groups? This is particularly important given that rent stabilization is not a means-tested program. Answers to this question sheds light on the on the ex-post incidence as well as the equity, efficiency, and

potential misallocation of rent stabilization.

Four different types of sub-groups are considered. First, I distinguish between white tenants versus non-white tenants. Secondly, I differentiate by skill level, where low-skilled tenants are those who do not have a college degree while high-skilled tenants hold at least a college degree.[50] Moreover, I follow Diamond et al. (2019) and divide the sample by age and tenure duration.[51]

The IV estimation results for different sub-groups are documented in Table 1.7. Consistent with the main IV results in Table 1.6, there are no significant effects across different sub-groups on either labor force participation (panel A) or hourly wage rate (panel C). Again, the discussion is therefore focused on unemployment (panel B).

I first distinguish between whether tenants are white or nonwhite (Hispanic, African-American, Asian, etc.) On the one hand, white tenants have a relatively lower unemployment rate (4.5 percent) on average compared with nonwhite tenants (6.7 percent). On the other hand, the negative effect of rent stabilization on unemployment is only significant among white tenants, which is statistically significant at the 1 percent level. In contrast, no significant effect is found among non-white tenants.

Second, I am particularly interested in the distinction by skill (or educational attainment). Surprisingly, the estimated effect is not significant among low-skilled

[50] About 18 percent of all tenants in the analytic sample hold post-graduate degrees.

[51] Diamond et al. (2019) use 4 years of tenure of residence to distinguish between short tenants versus long tenants. I follow their duration cutoff, which is also around the mean of the sample in my paper.

tenants but is only significant among high-skilled tenants (at the 1 percent level) with larger effects. Why are effects only significant among high-skilled tenants who are commonly found to have shorter unemployment spells compared with low-skilled tenants? I offer one novel explanation that is discussed in the next section - the policy awareness.

I then distinguish tenants by age and tenure duration. The effect is only significant among tenants whose ages are between 41 and 55 while no significant effect is found among tenants whose ages are between 26 and 40. Lastly, I do not have find any significant patterns regarding short-term versus long-term tenants.

Sorted Effects Method. Furthermore, the sorted effects method developed in Chernozhukov et al. (2018) is applied to understand the heterogenous effect beyond the average. Since the sorted effects method works with nonlinear models, I first apply this method to an ordinary probit model without using IV. The average marginal effect is about 3.6 percentage point, which is displayed by the black line in the left panel of Figure 1.8).[52] In contrast, the sorted partial effect varies strongly from 0 to about 9 percentage point, which is given as the blue line in the left panel of Figure 1.8. It does not coincide with the average effect of 3.6 percentage point.

Ideally, I am more interested in applying the sorted effects method to the IV estimation, so that the estimation results can be interpreted as causation directly. However, the difficulty lies in the fact that the sorted effects method applies to non-

[52]This has been mentioned previously in Table 1.15.

linear models while the previous IV estimation is based on two-stage least square estimation in which both first stage and second stage estimations are based on linear models. When the outcome variable is continuous, one can apply the control function method in which the first stage model can be nonlinear. However, for discrete outcome variable, a traditional plug-in approach that would replace the rent stabilization dummy with probit fitted values in the second-stage probit model, is nonparametrically unidentified and inconsistent (Chesher, 2003; Imbens and Newey, 2009; Wooldridge, 2015).

An alternative approach is currently applied by adding residuals from the first-stage probit model to the second-stage probit model (Terza et al., 2008). This might provide an accurate correction for "small" amounts of endogeneity (Wooldridge, 2015). I admit that this approach may not fully remove all the endogeneity concerns. Thus, the result is mainly for illustration purposes and the interpretation needs caution. Nevertheless, the sorted partial effect varies strong from 0 to about 12 percentage point, shown by the blue line in the right panel of Figure 1.8). It does not coincide with the average effect of 7.5 percentage point either, shown by the black line in the right panel of Figure 1.8.

1.7 Does Policy Awareness Matter?

So far, the empirical analysis has showed that rent stabilization *causes* higher unemployment while has no significant impact on labor force participation or hourly wage rate. This section moves beyond the classical assumption that program participants are always aware of their participation status. Instead, I propose a novel strategy that exploits the information set of rent-stabilized tenants and ask whether policy awareness matters.

Theoretically, being aware or not affects tenants' forecasting on future rent increases. For example, rent-stabilized tenants who are aware of such policy benefits would have a more stable expectation with lower annual rent increases than both rent-stabilized tenants who are unaware and tenants in private market-rate units in superstar cities like New York City. Being aware would further affect other types of forward-looking decision making. It may also make effects more salient based on insights from behavioral economics.

1.7.1 Not Everyone Knows!

A unique feature of the information in the 2002 and 2005 waves of the NYCHVS is that householders are asked to self-report their rent regulation status in addition to the available legal rent regulation status verified from administrative sources.[53] One might expect that most rent-stabilized tenants are aware of their good fortune,

[53]Unfortunately, such information has been unavailable after 2005.

that is, those who live in rent-stabilized units should know that their units are indeed rent-stabilized. However, the simple summary statistics provided in Table 1.8 show that clearly not every knows!

The first column in Table 1.8 lists the possible answers a tenant can report: (1) under rent control, (2) under rent stabilization, (3) private market-rate unit (not regulated), (4) do not know, and (5) do not report. The second and third columns report the two possible legal regulation situations in the data: either private market-rate (column (2)) or rent-stabilized (column (3)). This generates a 5×2 matrix where each cell represents a specific pair of self-reported and legal rent regulation status.

Let's first focus on the third column, where all participants live in a rent-stabilized unit. However, not all rent-stabilized renters know about their good fortune. It may be more surprising than one expects: approximately 34 percent are properly aware (8.60 percent choosing (1) "Under Rent Control" and 25.55 percent choosing (2) "Under Rent Stabilization"). In other words, these rent-stabilized tenants think their units are either rent-stabilized or rent-controlled.[54] In contrast, there are more than 24 percent of rent-stabilized tenants who think their housing units are private market-rate units that are not regulated at all. Those tenants are clearly misinformed. There are about 20 percent of rent-stabilized tenants who simply do not know their

[54]Given that housing units without rent control could still be rent stabilized, I think it may be too strict to require tenants to be able to tell the distinction between rent control and rent stabilization. I also argue that the distinction between rent-controlled and rent-stabilized is not very meaningful.

regulation status.[55]

Why would so many rent-stabilized tenants not be able to correctly specify their own housing regulation status? One possibility may be that the question is not well-posed during the survey. One good way to verify this concern is to look at the second column, which shows all the tenants who live in private market-rate units. Most private market-rate renters (around 60 percent) have chosen the correct answer. In contrast, there are less than 5 percent of private market-rate renters who are misinformed (1.79 percent choosing (1) "Under Rent Control" and 3.48 percent choosing (2) "Under Rent Stabilization"). This means that they believe their apartments are rent regulated. The misinformation rate among private market-rate tenants is much lower than the rate among rent-stabilized tenants. Therefore, I am confident that the survey question contains useful information.

Distinguishing between legal stabilization status and self-reported status provides a novel strategy for investigating the role of policy awareness by dividing rent-stabilized tenants into two groups: aware and misinformed. The two groups are defined as the follows:

1. Aware tenants: rent-stabilized tenants who choose either "Under Rent Control" or "Under Rent Stabilization" as the self-reported regulation status.

2. Misinformed tenants: rent-stabilized tenants who choose "Private Market-Rate" as the self-reported regulation status - they clearly have the wrong answer and

[55]Twenty percent of rent-stabilized tenants choose not to report any answer for this question. They are excluded in the analysis.

are misinformed.[56]

Descriptive Analysis. Before moving to the empirical analysis, let's first compare the aware (and the misinformed) rent-stabilized tenants with the private market renters in terms of their socioeconomic and demographic characteristics. The previous section demonstrated how rent-stabilized tenants are observationally similar to tenants in private market-rate units without making such policy awareness distinction.

As shown in Table 1.18, when comparing the aware group (column (1)) to the private market-rate renters (column (3)), these two groups are considerably different from one another.[57] Perhaps most important, the aware group has considerably higher educational attainment than private market-rate renters. More than 50 percent of aware and rent-stabilized tenants have at least a college degree, as compared with less than 40 percent of private market-rate renters. Similarly, around 10 percent of aware rent-stabilized tenants are high school dropouts and 16 percent are high school graduates. Both of these numbers are larger among private market-rate renters (15 and 24 percent, respectively). In terms of labor market outcomes, the aware renters

[56] An alternative definition could include a set of unaware rent-stabilized tenants, who either choose "Private Market-Rate" or "Don't Know" as the self-reported regulation status. However, one may express concern regarding the underlying meaning and interpretation of "Don't Know.". For example, if a tenant knows that the housing unit is rent-regulated but is unsure whether the housing unit belongs to the rent control or rent stabilization specification, such a tenant may actually choose "Don't Know." I express gratitude towards Katherine O'Regan for suggesting this perspective. Such "Don't Know" tenants should actually be classified in the aware group instead of the unaware group. Defining misinformed tenants avoids any such potential ambiguity. The empirical results are quantitatively similar but less precisely estimated when the *unaware* tenant instead definition is used instead of *misinformed* tenant definition. Results are available upon request.

[57] An overall comparison among the three groups of tenants is documented in Table 1.16

have both a higher labor force participation rate and higher unemployment rate - both of which are statistically significant - than private market-rate renters. The aware renters even have a relatively higher wage income (and wage rate) as well as nonlabor income than private market-rate renters.

In contrast, the misinformed group (Table 1.19) is significantly different when compared with private market-rate renters. In particular, the misinformed group expresses worse educational attainment. Less 35 percent of misinformed rent-stabilized tenants hold college degree or above, which is 5 percent lower than private market-rate renters and 15 percent lower than the aware group). Almost 45 percent of unaware rent-stabilized tenants are only high school graduates or dropouts. The unaware group also expresses lower income. However, there is no significant difference in terms of labor force participation and only a weakly larger unemployment rate among the misinformed compared to private market-rate renters.

1.7.2 Policy Awareness and Unemployment

This section investigates whether policy awareness matters for the impact of rent stabilization on unemployment. Since being aware is likely to be endogenous, I propose an additional instrumental variable strategy by exploiting the local awareness knowledge differences at the sub-borough level to reach causation.

Preliminary Analysis. This analysis is similar to the empirical model I have previously adopted, where the only difference is that the rent stabilized dummy (RS_{ijst})

is decomposed into two groups, aware and misinformed:

$$Y_{ijst} = \alpha + \beta^A \text{Aware}_{ijst} + \beta^B \text{Misinformed}_{ijst} + \eta X_{ist} + \psi V_{jst} + \lambda Z_s + \pi W_t + \epsilon_{ijst} \quad (1.6)$$

Y_{ijst} is the outcome variable of interest and I focus on unemployment in this part. Aware_{ijst} equals 1 when a householder i living in a rent-stabilized unit j in sub-borough s and year t is aware, and it equals 0 when either a householder is misinformed or lives in a private market-rate unit. Similarly, $\text{Misinformed}_{ijst}$ equals 1 when a householder i living in a rent-stabilized unit j in sub-borough s and year t is misinformed, and equals 0 otherwise. Therefore, I am essentially comparing both the aware and misinformed groups with the private market-rate renters, or when $\text{Aware}_{ijst} = 0$ and $\text{Misinformed}_{ijst} = 0$. The coefficients of interest are both β^A and β^B. The other variables are defined in the same way as previously demonstrated.

There is a high possibility that, besides rent stabilization, policy awarenesss is also endogenous, particularly given the descriptive analysis provided earlier.

Instrumental Variable Strategy. To deal with self-selection, one must recognize the two endogenous variables here: tenants being either aware or misinformed conditional on living in rent-stabilized units. This requires at least two instrument variables to exogenously shift both margins. One option is to use a previously implemented IV, for example, $\text{IV}^{stab}_{b,t-1}$, which deals with sorting into rent-stabilized units. However, additional instruments are needed that exogenously shift the policy awareness. Inspiration comes from Chetty et al. (2013) who proxy for the Earned Income Tax

Credit (EITC) knowledge with the fraction of individuals who manipulate reported self-employment income to maximize their credit refund. Such knowledge varies significantly across areas. Similarly, I use the variation of rent regulation knowledge at the local neighborhood level (sub-borough level). More specifically, I use the share of rent-stabilized tenants who are aware ($\text{IV}_{sb,t}^{aware}$), misinformed ($\text{IV}_{sb,t}^{mis}$), or don't know ($\text{IV}_{sb,t}^{dk}$) in each sub-borough as the additional instruments

It is worth mentioning that all $\text{IV}_{sb,t}^{aware}$, $\text{IV}_{sb,t}^{mis}$ and $\text{IV}_{sb,t}^{dk}$ are at the sub-borough level, which is much smaller and local than the borough level. In addition, they are at the same year as when labor market outcomes are measured because self-selection into rent-stabilized units has been taken care of by the inclusion of $\text{IV}_{b,t-1}^{stab}$.

I first briefly discuss the validity of the proposed instruments, which follows closely with the discussion of the validity of $\text{IV}_{b,t-1}^{stab}$. On the one hand, the first-stage result relies on meaningful geographic variation that affects individual tenant's housing regulation knowledge - policy awareness. As shown in Figure 1.9 (and Figures 1.14, 1.15), there is significant variation at the sub-borough level with respect to local knowledge about rent stabilization. On the other hand, in terms of exogeneity, I discuss the share of rent-stabilized tenants who are aware ($\text{IV}_{sb,t}^{aware}$) for illustration. One may expect that being aware is highly correlated with educational attainment, therefore Manhattan should have the highest spatial concentration of such local awareness. However, Figure 1.9 clearly shows that there are other non-Manhattan sub-boroughs that have a large share of local regulation awareness. Moreover, there are significant differences

even between two adjacent sub-boroughs. Last, I control for education, ethnicity, and income level at the sub-borough level in addition to the sub-borough and year fixed effects to tease out other possible confounding factors at the sub-borough level.

Estimation Results. This section discusses results from both OLS and IV estimates. First, OLS estimates are reported as the baseline analysis in Table 1.9. In the first column, the OLS result suggests that living in a rent-stabilized unit and being aware is associated with a 1.9 percentage point higher likelihood of being unemployed when compared with private market renters. This is significant at the 5 percent level. In contrast, living in a rent-stabilized unit and being misinformed is associated with a 1.5 percentage point higher likelihood of being unemployed when compared with private market renters. This finding is not statistically significant. Both probit and logit models yield similar results that are aligned with the OLS estimates. However, since both the aware group and misinformed group are observationally different when compared with private renters, respectively, the estimated coefficients are unlikely to be causal.

I then turn to the IV estimations. The first stage result is shown in Table 1.10. Regarding IV (1), the sign of the coefficients is consistent with expectation. The first column marks the endogenous variable as being aware. Here, if a sub-borough has more tenants who are aware on average, then individual rent-stabilized tenants are more likely to be aware. In contrast, if a sub-borough has more tenants who are misinformed on average, then individual rent-stabilized tenants are less likely to be

aware. The opposite is true for misinformed tenants in the second column. All of these estimates are conditional on the sub-borough characteristics and fixed effect. One must also recognize that the original IV implemented for rent stabilization - the fraction of vacant-for-rent and rent-stabilized at the borough level - is statistically significant for being aware while insignificant for being misinformed. The share of rent-stabilized tenants who are misinformed or don't know are used as an alternative set of instruments - IV (2).[58] It is interesting to observe that when a sub-borough has a larger share of rent-stabilized tenants who report "Don't Know", then individual tenants are more likely to be aware and less likely to be unaware. This suggests that "Don't Know" may contain some hidden information.

The IV results for unemployment are shown in Table 1.11. The negative impact of rent stabilization on unemployment is only significant among the aware tenants. By removing unobserved confounding factors at the individual level, the estimated causal effect is much larger than the OLS estimates. This effect is probably not surprising, given that the aware tenants have a much better educational attainment, are more likely to be born in New York City, and are more likely to be white. By removing these socioeconomic premium, the actual causal effect becomes larger. However, there is no significant effect among the misinformed tenants when compared with renters in private market-rate units. In addition, the F-statistics are around 10. This further strengthens confidence in the results.

[58] Using all three IVs jointly yield similar results quantitatively. Results are available upon request.

Last, the results discussed here could also offer a rationale for one of the puzzling findings in the previous section: why is the negative effect of rent stabilization on unemployment is only significant among high-skilled and white tenant but not among socioeconomically disadvantaged tenants (such as low-skilled or ethnically minority tenants)? One likely explanation is because the policy awareness is also concentrated among high-skilled and white tenants who live in rent-stabilized units. Such policy awareness would further strengthens the underlying mechanisms - both the insurance and mobility channels. There are strong evidence showing that quality-adjusted rent discounts are much larger for tenants who are aware, and the aware tenants also have much longer residential tenure compared with both misinformed tenants and private market-rate tenants.

1.8 Discussions

1.8.1 Rent Discount and Mobility as Alternative Mechanisms

The previous analyses are based on a binary dummy variable that indicates whether a tenant lives in a rent-stabilized unit or private market-rate unit. While this definition captures the extensive margin, it does not succeed in regards to the intensity margin. This is important because rent-stabilized tenants may have differ-

ent levels of rent discount benefits. Therefore, rent discount is used as an alternative measure to test the robustness of the results. It is reasonable to expect that a larger rent discount is associated with a larger effect on unemployment. Moreover, I add tenure of residence, which is measured by how many years a tenant has lived in the current dwelling as an additional covariate. This may shed light on the possible mechanisms between the rent discount and "lock-in" (mobility) channels. The results are reported in Table 1.12, wherein the outcome variable is unemployment dummy.

The first column reproduces the result using the rent stabilization dummy from the previous section. The second column, however, uses the continuous rent discount measure (Jiang et al., 2019) instead of the rent stabilization dummy. The effect is still statistically significant at the 1 percent level. More specifically, a $1,000 monthly rent discount resulting from rent stabilization is associated with a 1.9 percentage point higher likelihood of being unemployed on average. The third column adds tenure of residence on top of the rent discount variable. Even though the coefficient in front of rent discount has been slightly absorbed, it remains significant at the 5 percent level. In contrast, even though the coefficient in front of tenure of residence is positive, it is not significant. One might be concerned that by having a single tenure of residence variable, a strong assumption that every additional year of living in the same unit yields the same effect is imposed. As an alternative specification, I discretize the mobility measure into different categories. As shown in the fourth column, the pattern remains the same. Namely, a higher rent discount is associated

with a higher likelihood of being unemployed (at the 1 percent level) while mobility measures are not significant.

The results might be suggestive in terms of the relative importance of the two different mechanisms. However, the interpretation demands caution because (1) both rent discount and mobility are truly endogenous variables, which may also depend on unobservable heterogeneity; and (2) the analytic sample only includes tenants who currently live in New York City, so people who move out of New York City to other states cannot be tracked.

1.8.2 Sensitivity Analysis

Alternative Instruments for Rent Stabilization. As a robustness check, I also adopt alternative instruments by using the numerators and denominators separately associated with $\text{IV}_{b,t}^{stab}$ and $\text{IV}_{b,t}^{mkt}$. This also allows for an over-identifying restriction test given that the number of IVs are larger than the number of endogenous variables. Therefore, I therefore use the total number of vacant-for-rent units, as well as vacant-for-rent units being rent-stabilized (or market-rate), as instruments, respectively. The same logic still applies. The difference is that I now rely on the change of absolute number of different types of housing units directly, while previous IVs use the relative change in shares of different types.

The first stage results using the alternative instruments are documented in Table 1.20. The direction of each IV is still as expected. For example, the first column

shows a larger total number of vacant-for-rent and rent-stabilized units leads to a higher probability of living in a rent-stabilized unit. More specifically, an additional 1,000 vacant-for-rent and rent-stabilized units leads to a 2.8 percentage point higher likelihood to live in a rent-stabilized unit. The opposite is true for the total number of vacant-for-rent and market-rate units in the second and third columns.

The second stage results for unemployment are documented in Table 1.21[59]. The qualitative story holds exactly the same as for the previous IV results: rent stabilization does significantly lead to higher unemployment. Using alternative IVs leads to a relatively smaller estimated coefficient - about 1-2 percentage point smaller in the absolute number. Moreover, the analyses do not reject the over-identifying restriction tests. Overall, the estimates are robust and provide convincing causal interpretation.

Missing Values in Policy Awareness. Lastly, I consider the implications of excluding those observations who have missing values in the policy awareness variable as additional robustness check. As shown in Table 1.8, there are about 20% of tenants have chosen "not reported" in answering to the question about self-reported rent regulation status. Such tenants with missing values in the policy awareness variable are excluded in the analytic sample in the previous section. However, one might be concerned that such missing values are not random, which might create a sample selection problem. Hence, I also include these tenants as an additional "missing" category, in addition to the original "aware" and "misinformed" categories. Descriptive

[59]I do not report result on labor force participation and wage rate to save space, given their insignificant results that are consistent with previous findings. Results available upon request.

analysis shows that tenants in this "missing" category are observationally more similar to renters who are private market renters, while much less similar to renters who are misinformed. [60] Furthermore, empirical results from both linear regression model and instrumental variable estimations are shown in Table 1.22. [61] It is confirmed that the unemployment effect of rent regulation mainly comes from tenants who are aware, even when the missing category is added.

1.9 Conclusion

This paper takes a first step in understanding the relationship between rent regulation and tenant labor market outcomes and emphasized the role of policy awareness. I find that rent stabilization causes higher unemployment for rent-stabilized tenants comparing with tenants in private market-rate units. Effects are more significant among high-skilled, white tenants. I take novel usage of a unique data set on rent stabilization in New York City. To deal with self-selection into rent-stabilized units, I propose an IV strategy. I leverage variation in the availability of rent-stabilized units across New York boroughs over time.

Moreover, this paper moves beyond traditional program evaluation assumptions and asks whether policy awareness matters. The unique data feature allows for dif-

[60] Descriptive summary statistics are available upon request.

[61] There are now three endogenous variables: rent-stabilized and aware, rent-stabilized and misinformed, and rent-stabilized and missing awareness. So at least three instrumental variables are needed to achieve identification. In addition to the one instrumental variable used for rent stabilization or not, I use both the share of rent-stabilized tenants who are aware, and share of rent-stabilized tenants who are misinformed at sub-borough level.

ferentiating between legal regulation status versus tenants' perception. The rent-stabilized tenants are classified into two groups: aware and misinformed. Aware tenants know their housing units are regulated. Misinformed tenants live in rent-stabilized units but believe their units are not regulated. Surprisingly, about 25 percent of rent-stabilized tenants are misinformed about their good fortune. Furthermore, the negative impact of rent stabilization on unemployment is only significant among tenants who are aware when local housing regulation knowledge is used as additional instruments to reach causation.

This paper sheds light on important policy implications, particularly in the presence of renewed legislative momentum at the local municipal level to impose new rent control measures. The findings are particularly relevant given the recent rent regulation reform in Oregon, New York, and California since 2019. Rent stabilization is almost like a universal in-kind transfer program.[62] It is not means-tested and poorly targeted. This paper provides novel causal evidence that rent stabilization causes higher unemployment, which has important aggregate implications. Given the data limitation, it is hard to further tell whether higher unemployment could be owning to either eased liquidity constraint or distorted job search behavior, both of which are fundamental in the discussion of optimal unemployment insurance literature (Moffitt, 2014).

The results still offer implications for potential policy improvements. On the

[62]This may also be related to the recent discussion on the universal basic income (UBI) program (Kearney and Mogstad, 2019) in the United States.

one hand, if the higher unemployment effect is due to eased liquidity constraint - allows unemployed rent-stabilized tenants to search longer for better jobs, then such unemployment effect might be socially desirable. However, from a redistribution perspective, who deserves better access to eased liquidity constraint? The current policy seems to be *inequitable* because the effect is only significant among high-skilled, white tenants. This can also be interpreted as additional type of misallocation cost (Glaeser, 1996; Glaeser and Luttmer, 2003) broadly.

On the other hand, if the higher unemployment effect is because of distorted job search behavior, such effect is thus *inefficient* and socially undesirable.[63] If this is the case, policy makers could learn from the enforcement of minimum job search effort, which is commonly considered in both unemployment insurance and other welfare programs, to counteract such inefficiency.

Overall, housing subsidy programs (such as the housing voucher) which target lower income families might be much better than rent control or rent stabilization policies, because housing subsidy programs would have more favorable distributional effects than rent stabilization. The recent report "a roadmap to reducing child poverty" by National Academies of Sciences, Engineering, and Medicine[64] has projected that expanding the number of housing vouchers would result in major reductions in the U.S. poverty rate. Given that current housing subsidies are limited, millions of low

[63]Another possibility maybe that rent stabilization induces tenants to quit their jobs more frequently than optimal (Topel and Ward, 1992).

[64]National Academies of Sciences, Engineering, and Medicine 2019. A Roadmap to Reducing Child Poverty. Washington, DC: The National Academies Press. https://doi.org/10.17226/25246.

income families cannot even get a voucher, and many families with vouchers are discriminated by landlords, it seems expanding housing subsidy programs such as the housing voucher with better enforcement would be more desirable than bringing back rent regulation.

In addition to the efficiency and equity consideration, this paper also offers novel evidence that suggests policy awareness matters. The effects on unemployment are only significant among tenants who are aware, which may also suggest a new direction for policy improvement as well as future research - what would happen if landlords are not allowed to inform tenants about the rent stabilization status?

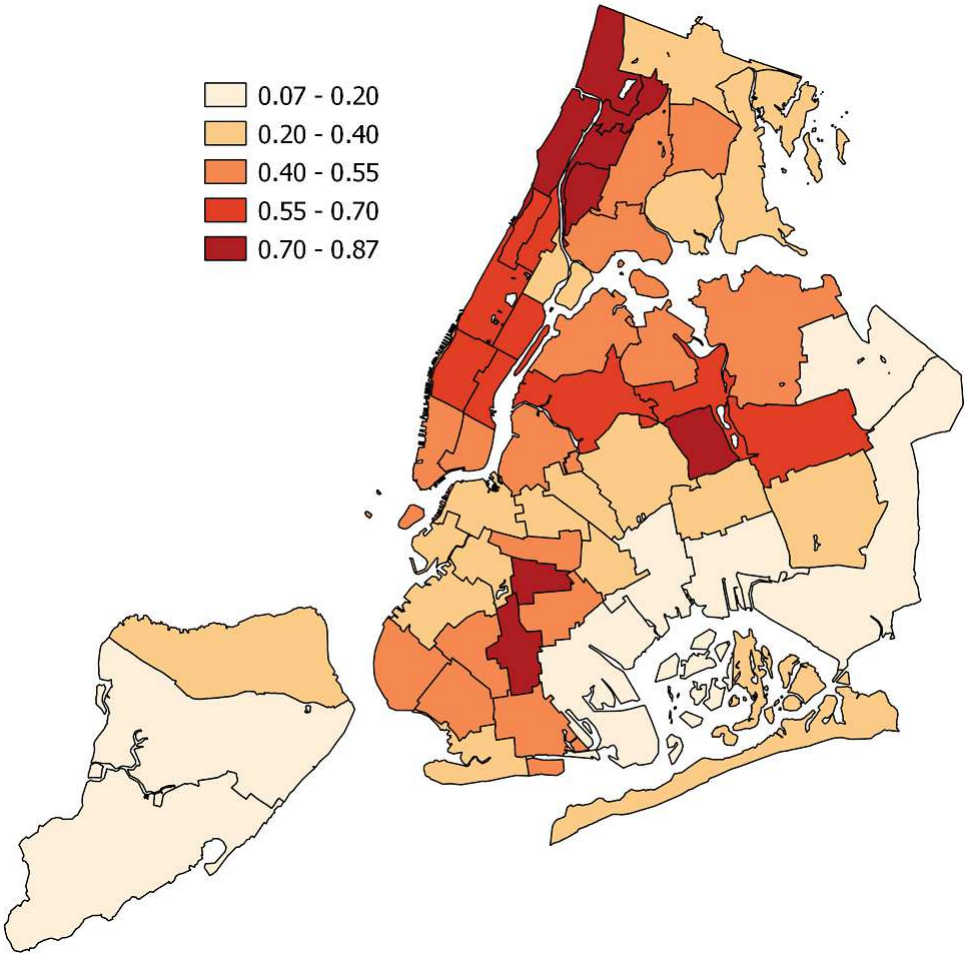
There are multiple avenues for future research. For example, since the data are not longitudinal, I am unable to investigate the effect of rent stabilization on unemployment spell duration and re-employment wage. Both outcome variables are important measures to better understand the implication on unemployment. Moreover, I am not able to observe households when they move to other places or outside New York City. This would require a similar panel data structure as used in (Diamond et al., 2019). Having better data with more measures about job search and labor market behaviors, as well as longitudinal structure, would allow for deeper understanding of the underlying mechanisms to complement the findings of this paper.

Another important extension would be to investigate the interaction between rent regulation and neighborhood effects (Chetty et al., 2016; Chyn, 2018). Since rent stabilization leads to longer residential stability, it could potentially have different effects

for not only adults contemporaneously but also for children in the long run, depending on different neighborhood qualities. This would be important for understanding the policy consequences on human capital development and intergenerational mobility. These are left for future studies.

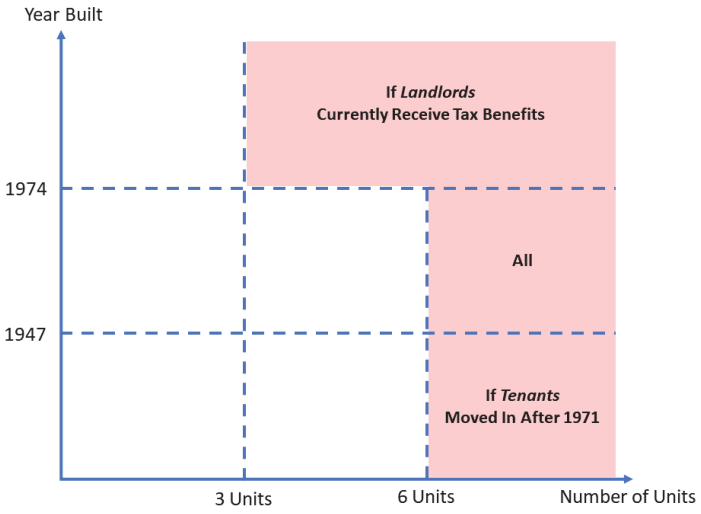
1.10 Figures and Tables

Figure 1.1: Spatial Concentration of Rent-Stabilized Units in NYC



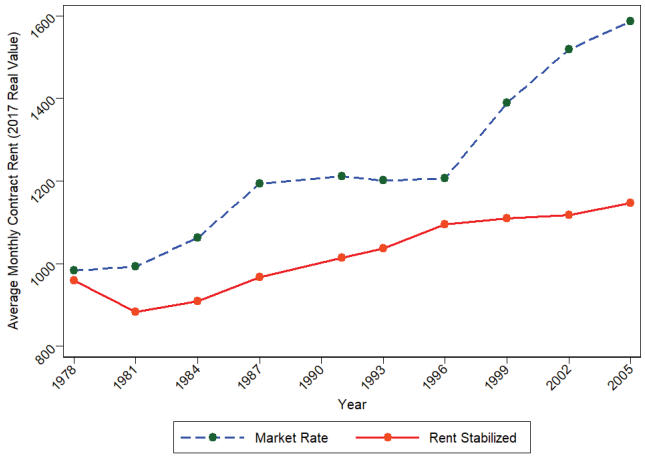
Notes: Author's calculation based on pooled 2002 and 2005 waves of New York City Housing and Vacancy Survey (NYCHVS). Darker color represents higher concentration of rent-stabilized units in the rental market in a sub-borough.

Figure 1.2: Eligibility of Rent Stabilization



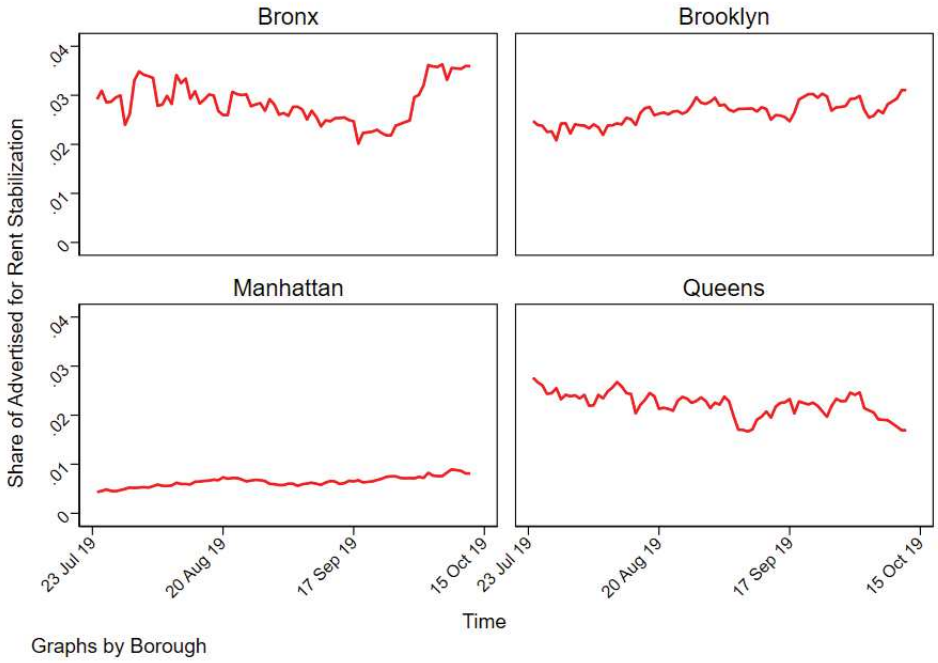
Notes: Rent stabilization in New York City generally applies to three types of apartments in buildings of: (a) six or more units built between February 1, 1947 and January 1, 1974; (b) six or more units built before February 1, 1947 and tenants who moved in after June 30, 1971; (c) three or more apartments constructed or extensively renovated since 1974 with special tax benefits such as J-51, 421a, or other programs. This only applies to the period when tax abatement is effective, which usually lasts for ten to twenty years. This figure only considers rent-stabilized units that are not subject to de-regulation.

Figure 1.3: Average Rent Comparison in New York City



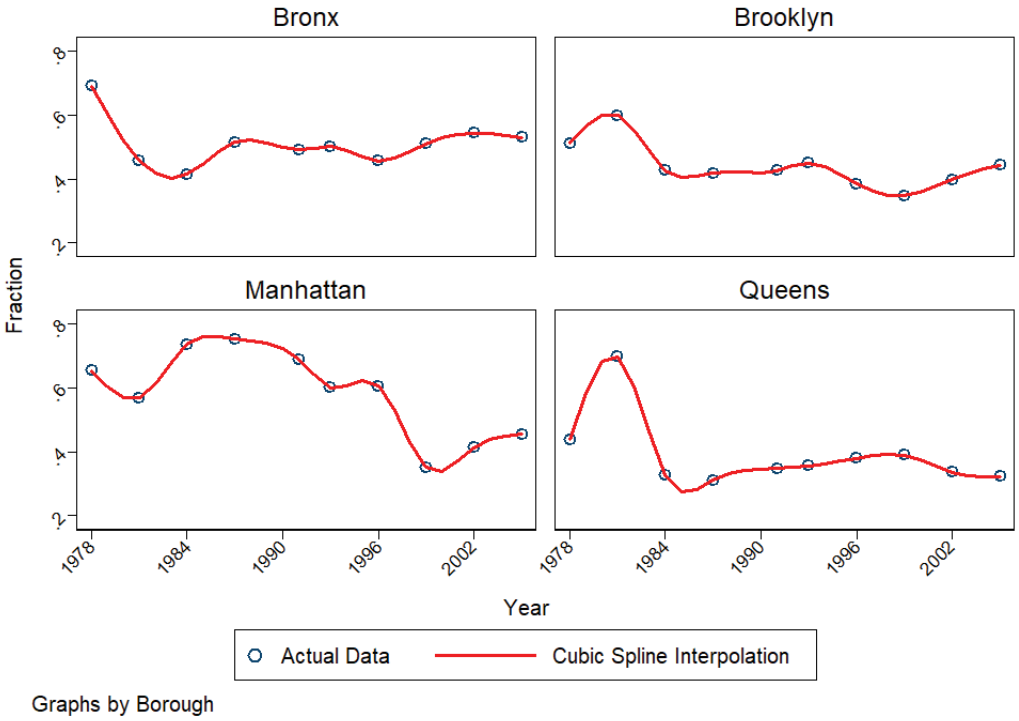
Notes: Data is obtained from 1978 - 2005 waves of the New York City Housing and Vacancy Survey (NYCHVS). Rent is measured as average monthly contract rent in 2017 $1000 real value. The solid line denotes the average monthly contract rent for rent-stabilized units while the dashed line denotes the one for private market-rate units.

Figure 1.4: Share of Online Postings with Rent Stabilization Advertisement



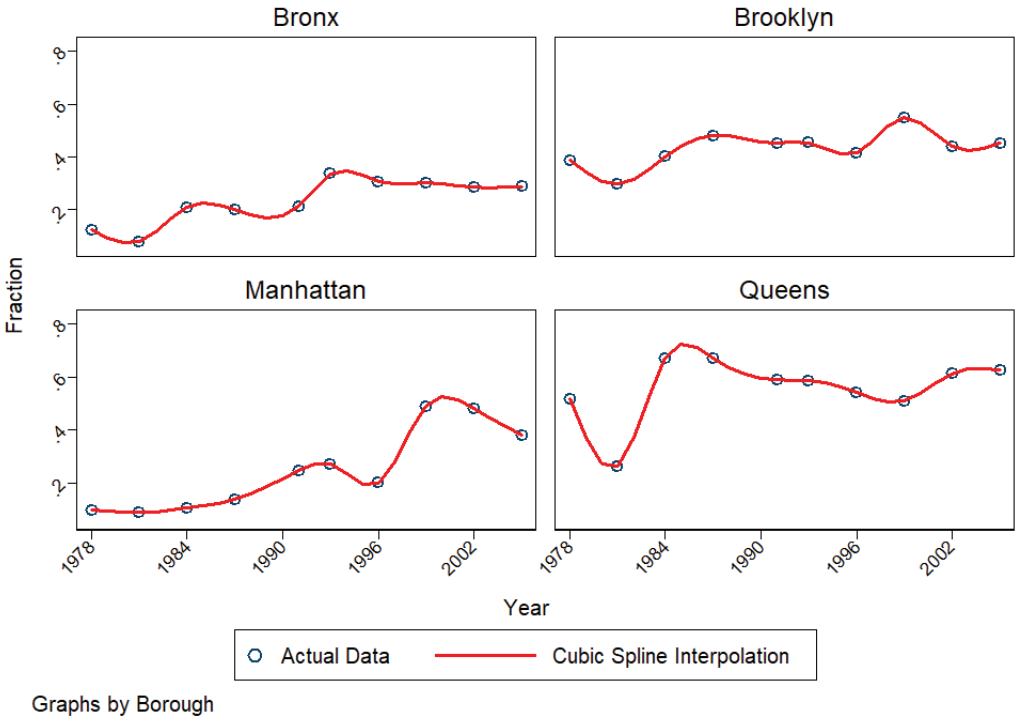
Notes: Data is obtained by author's manual collection from `StreetEasy` website (`https://streeteasy.com`). An online posting is considered to have rent stabilization advertisement if any of these word appear in the description section: "rent stabilization", "rent stabilized", "stabilized", etc. Historical New York City Housing and Vacancy Survey (NYCHVS) data suggests that at least 25% of total vacant-for-rent units are rent-stabilized. This suggests only a small minority of actual rent-stabilized, vacant apartments are advertised for their stabilization status.

Figure 1.5: The Variation of $\text{IV}^{stab}_{b,t-1}$ over Time by Borough



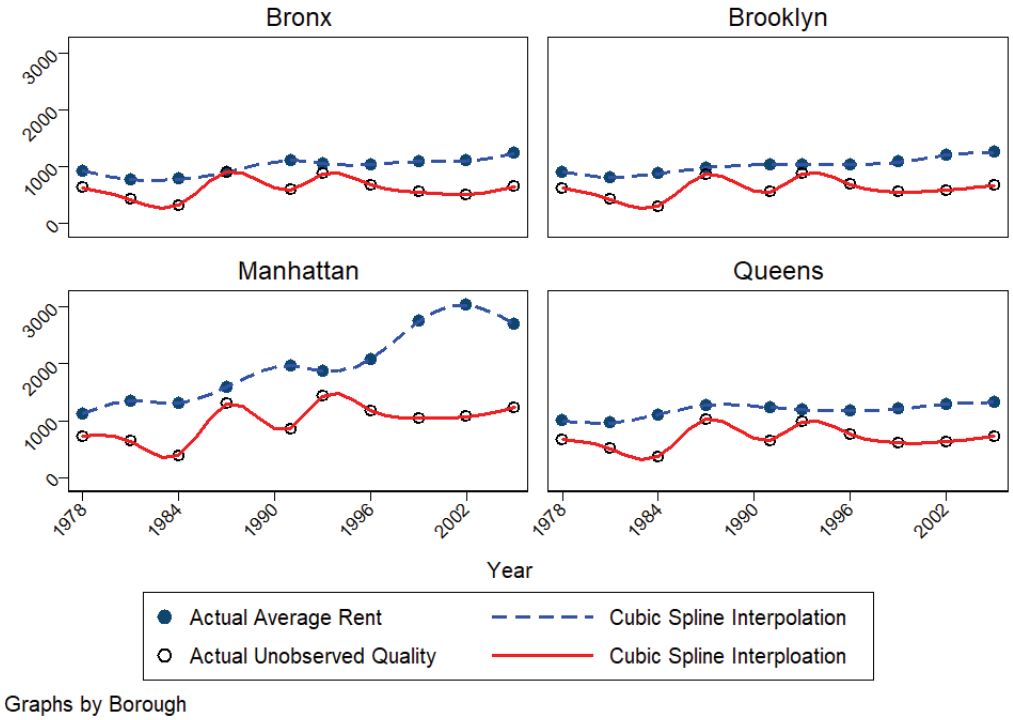
Notes: The instrument $\text{IV}^{stab}_{b,t-1}$ is defined as the ratio of the total number of vacant-for-rent units that are rent-stabilized $N^{stab}_{b,t-1}$ and the total number of vacant-for-rent units $N^{all}_{b,t-1}$. Blue dots are directly obtained and calculated from 1978-2005 waves of the New York City Housing and Vacancy Survey (NYCHVS), while the red lines represent fitted value based on cubit spline interpolation. Staten-Island has fairly small share of rent stabilized rental units and is available upon request.

Figure 1.6: The Variation of $\text{IV}_{b,t-1}^{mkt}$ over Time by Borough



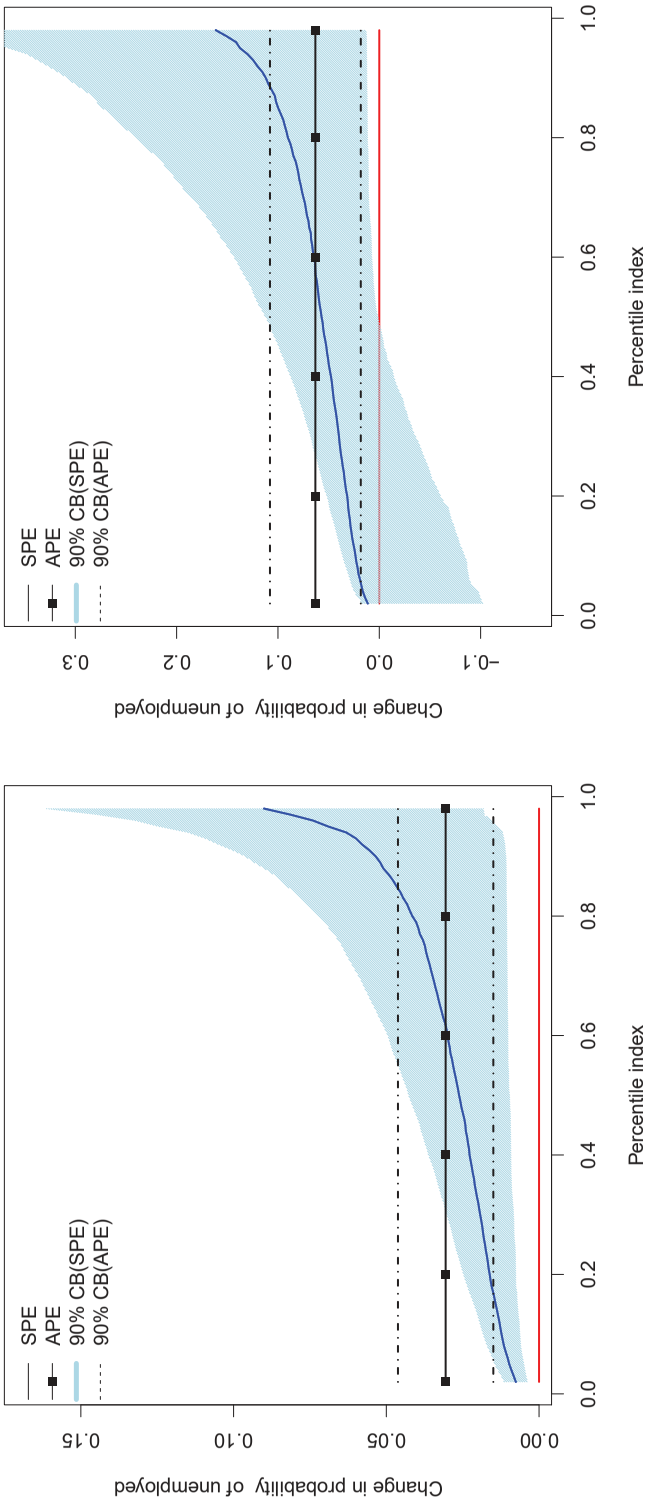
Notes: The instrument $\text{IV}_{b,t-1}^{mkt}$ is defined as the ratio of the total number of vacant-for-rent units that are private market-rate $N_{b,t-1}^{mkt}$ and the total number of vacant-for-rent units $N_{b,t-1}^{all}$. Blue dots are directly obtained and calculated from 1978-2005 waves of the New York City Housing and Vacancy Survey (NYCHVS), while the red lines represent fitted value based on cubit spline interpolation. Staten-Island has fairly small share of rent stabilized rental units and is available upon request.

Figure 1.7: The Unobserved Neighborhood Quality by Boroughs



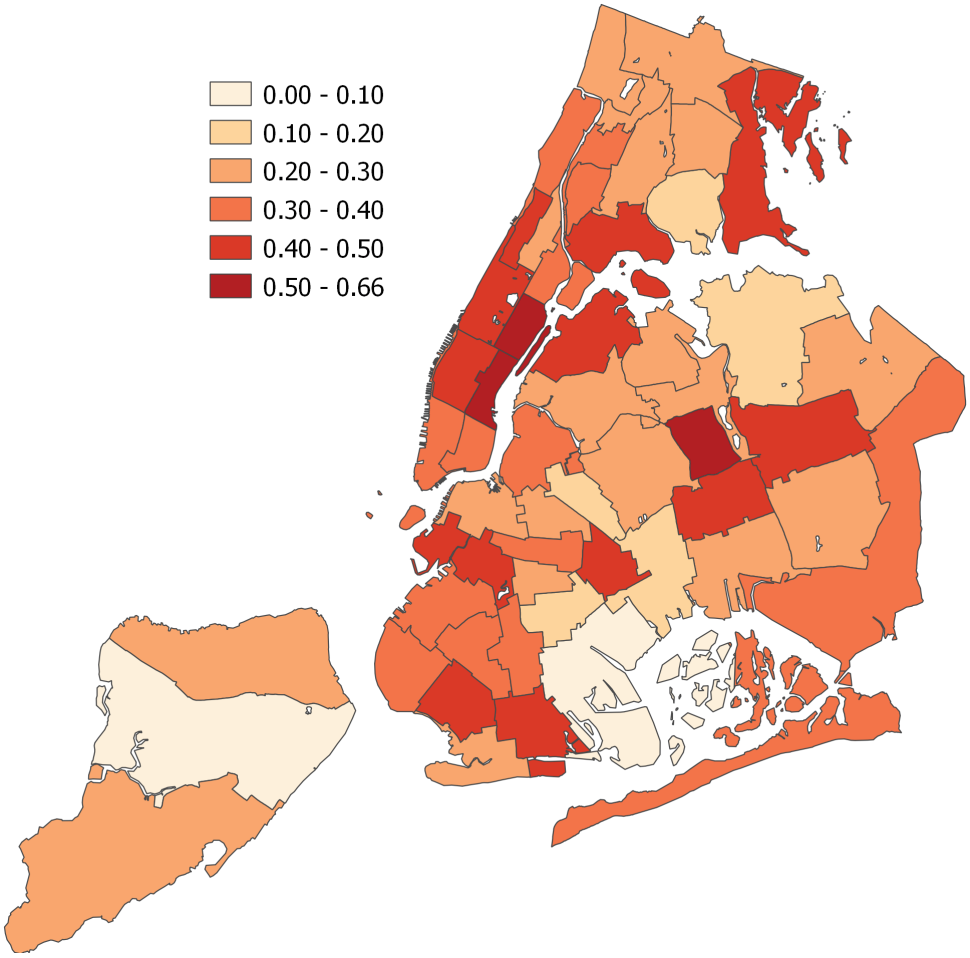
Notes: Y-axis is measured in terms of 2017 U.S. dollar value. Dots are directly obtained and calculated from 1978-2005 waves of the New York City Housing and Vacancy Survey (NYCHVS). Red solid lines represent fitted value for unobserved neighborhood quality based on cubit spline interpolation in each borough. Blue dashed lines represent fitted value for average rental price based on cubit spline interpolation in each borough. The result of Staten-Island is available upon request.

Figure 1.8: The Sorted Effects Method



Notes: Both figures apply the sorted effects method in Chernozhukov et al. (2018). In each figure, black solid line is the average marginal effect, dashed line is the confidence interval for the average marginal effect, blue solid line represents the collection of sorted effects, and light blue region is the confidence sets. In the left figure, I apply the sorted effects method to an ordinary probit model in Table 1.15. In the right figure I apply the sorted effects method to a probit model (the second stage) by adding residuals from a first-stage probit model using instruments for illustration purpose. The interpretation of the right panel needs caution.

Figure 1.9: Spatial Concentration of Policy Awareness (Aware)



Notes: Author's calculation based on pooled 2002 and 2005 waves of New York City Housing and Vacancy Survey (NYCHVS). Darker color represents higher concentration of rent-stabilized tenants who are aware among all rent-stabilized renters in a sub-borough.

Table 1.1: Summary Statistics by Rent Stabilization

	Rent-stabilized	N	Market-rate	N	Difference
			Panel A. Household Characteristics		
Currently In the Labor Force	0.89	5514	0.89	3908	0.00
Currently Unemployed	0.06	4904	0.05	3471	0.01**
HHH Wage Rate	29.02	4760	30.70	3333	-1.68***
High School Dropout	0.17	5514	0.15	3908	0.01
High School Graduate	0.23	5514	0.24	3908	-0.01
Some College No Degree	0.20	5514	0.21	3908	-0.01
College Degree and Above	0.40	5514	0.40	3908	0.00
Female	0.51	5514	0.45	3908	0.06***
Age	38.92	5514	38.19	3908	0.73***
White, Non-Hispanic	0.36	5514	0.41	3908	-0.06***
Black, Non-Hispanic	0.22	5514	0.22	3908	-0.00
Hispanic	0.31	5514	0.21	3908	0.10***
Asian	0.10	5514	0.14	3908	-0.04***
Married	0.35	5514	0.44	3908	-0.09***
Born in NYC	0.21	5514	0.24	3908	-0.03***
Born in U.S. (Non-NYC)	0.14	5514	0.13	3908	0.01
Born in Other Country	0.47	5514	0.45	3908	0.02**
Birth Place Missing	0.18	5514	0.18	3908	-0.00
Any Young Child in HH	0.16	5514	0.19	3908	-0.03***
Any Child in HH	0.36	5514	0.41	3908	-0.04***
Co-reside with Parents	0.04	5514	0.03	3908	0.00
HH Size	2.40	5514	2.69	3908	-0.29***
Any Non-relative in HH	0.11	5514	0.13	3908	-0.02***
HHH Salary Income, cond. work	52.56	4760	56.98	3333	-4.41***
HHH Non-Labor Income	2.95	5514	3.12	3908	-0.17
Other HH Mem. Tot. Inc.	20.04	5514	26.75	3908	-6.71***
			Panel B. Housing Characteristics		
Monthly Contract Rent	1.18	5514	1.49	3908	-0.31***
Monthly Gross Rent	1.27	5514	1.62	3908	-0.35***
Tenure of Residence	7.00	5514	5.05	3908	1.95***
Num. of Rooms	3.17	5514	3.91	3908	-0.73***
Num. of Bedrooms	1.34	5514	1.83	3908	-0.49***
Num. of Unit Problems	0.91	5514	0.46	3908	0.45***
Num. of Building Problems	0.26	5514	0.18	3908	0.07***

Notes: Data comes from the analytic sample of pooled 2002 and 2005 waves of the New York City Housing and Vacancy Survey (NYCHVS). All the monetary values are in 2017 real dollar. Wage rate is in $ and all the income-related variables are in $1000. "HH" stands for household and "HHH" stands for household head. Unit problems include toilet breakdown, heating equipment breakdown, presence of mice and rats, cracks or holes in interior walls, holes in floors, broken plaster or peeling paint on inside walls, water leakage, etc (see Table **??** for more details). Building problems include issues related to external walls, building windows, stairways, floors, etc (see Table **??** for more details). * p<0.10, ** p<0.05, *** p<0.010.

Table 1.2: Effect of Rent Stabilization on Monthly Contract Rent

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Panel A. Level of Monthly Contract Rent							
Rent Stabilized	-0.38***	-0.26***	-0.25***	-0.40***	-0.39***	-0.39***	-0.38***	-0.37***
	(0.067)	(0.059)	(0.053)	(0.085)	(0.079)	(0.076)	(0.075)	(0.074)
Stab × Duration of Tenure		-0.019***	-0.018***	-0.022***	-0.021***	-0.020***	-0.020***	-0.019***
		(0.0024)	(0.0024)	(0.0022)	(0.0019)	(0.0018)	(0.0018)	(0.0017)
$\overline{Y}$	1.297	1.297	1.297	1.297	1.297	1.297	1.297	1.297
N	9422	9422	9422	9422	9422	9422	9422	9422
R^2	0.425	0.444	0.457	0.548	0.569	0.574	0.581	0.623
	Panel B. Log of Monthly Contract Rent							
Rent Stabilized	-0.24***	-0.13***	-0.13***	-0.16***	-0.16***	-0.16***	-0.15***	-0.16***
	(0.033)	(0.030)	(0.028)	(0.046)	(0.041)	(0.038)	(0.038)	(0.037)
Stab × Duration of Tenure		-0.017***	-0.016***	-0.019***	-0.018***	-0.018***	-0.018***	-0.017***
		(0.0017)	(0.0017)	(0.0015)	(0.0014)	(0.0013)	(0.0013)	(0.0011)
$\overline{Y}$	0.164	0.164	0.164	0.164	0.164	0.164	0.164	0.164
N	9422	9422	9422	9422	9422	9422	9422	9422
R^2	0.347	0.376	0.385	0.463	0.492	0.500	0.507	0.552
Sub-borough and Year FE	✓	✓	✓	✓	✓	✓	✓	✓
Building Year Built			✓	✓	✓	✓	✓	✓
Structure Characteristics				✓	✓	✓	✓	✓
Leasing Type					✓	✓	✓	✓
Unit Amenity and Quality						✓	✓	✓
Building and Nbhd Quality							✓	✓
Borough Interaction								✓

Notes: Standard errors are in parentheses and clustered at sub-borough level. * p<0.10, ** p<0.05, *** p<0.010. Structure characteristics include: 1) number of rooms in the unit, 2) number of bedrooms in the unit, 3) number of units in the building, 4) number of stories in the building, 5) unit level, 6) whether apartment building owner lives in the same building. Leasing type includes: 1) the length of leasing, 2) whether electricity is included in rent, 3) whether gas is included in rent, 4) whether gas is included in rent, 5) whether other fuel is included in rent. Unit amenity and quality include: 1) fuel type, 2) additional fuel, 3) plumbing condition, 4) kitchen condition, 5) presence of mice and rats, 6) exterminator service, 7) cracks or holes in interior walls, 8) holes in floors, 9) broken plaster or peeling paint, 10) water leakage, 11) number of heat breakdown, 12) number of toilet breakdown. Building and neighborhood quality include: 1) elevator in building, 2) sidewalk to elevator/unit without using steps, 3) wheelchair access to street entry/elevator/unit entrance, 4) has issue of building's external wall/windows/stairways/floor, 5) overall rating of building quality, 6) boarded up structures in neighborhood, 7) self-rating of neighborhood quality. Full results containing coefficients of other variables are available upon request.

Table 1.3: Determinants of Rent Stabilization Eligibility

	A. LPM	B. Probit	C. Logit
Female	0.040***	0.040***	0.041***
	(0.011)	(0.011)	(0.011)
White, Non-Hispanic	-0.040*	-0.040*	-0.039*
	(0.022)	(0.020)	(0.021)
Black, Non-Hispanic	-0.019	-0.019	-0.022
	(0.027)	(0.028)	(0.028)
Age	0.010	0.0092	0.0096
	(0.0063)	(0.0062)	(0.0062)
Married	-0.013	-0.014	-0.013
	(0.016)	(0.016)	(0.016)
Co-reside with Parents	0.034	0.033	0.033
	(0.026)	(0.027)	(0.026)
Born in U.S. (Non-NYC)	-0.014	-0.013	-0.013
	(0.016)	(0.015)	(0.015)
Born in Other Country	0.020	0.021	0.022
	(0.017)	(0.017)	(0.017)
High School Dropout	0.0074	0.013	0.012
	(0.017)	(0.017)	(0.017)
High School Graduate	0.024	0.027	0.025
	(0.017)	(0.017)	(0.018)
Some College No Degree	0.012	0.015	0.013
	(0.017)	(0.017)	(0.017)
Spouse Has College Degree or Above	-0.032	-0.027	-0.027
	(0.021)	(0.019)	(0.019)
Any Child in HH	0.035	0.037*	0.036
	(0.021)	(0.021)	(0.022)
HH Size	-0.040***	-0.040***	-0.041***
	(0.0079)	(0.0077)	(0.0079)
Any Non-relative in HH	-0.035**	-0.035**	-0.034**
	(0.017)	(0.016)	(0.016)
HHH Total Income (Log)	-0.0011	-0.0010	-0.00099
	(0.0013)	(0.0013)	(0.0013)
Other HH Mem. Tot. Inc. (Log)	0.00028	0.00040	0.00036
	(0.0011)	(0.0010)	(0.0010)
Sub-borough and Year FE	✓	✓	✓

Notes: Standard errors are in parentheses and clustered at sub-borough level. * p<0.10, ** p<0.05, *** p<0.010. Column A is linear probability model (LPM). Column B and C are probit and logit models respectively, and the coefficients have been transformed to be marginal effects to ease interpretation and comparison to LPM.

Table 1.4: Effect of Rent Stabilization on Labor Market Outcomes (OLS)

	Panel A. Labor Force Participation						
Rent Stabilized	-0.0049	-0.0024	-0.0015	-0.0024	0.0067	0.0092	0.0099
	(0.0085)	(0.0082)	(0.0083)	(0.0084)	(0.012)	(0.0072)	(0.0072)
R^2	0.021	0.083	0.100	0.106	0.120	0.559	0.565
	Panel B. Unemployment						
Rent Stabilized	0.015***	0.015***	0.014***	0.015***	0.026***	0.026***	0.023***
	(0.0055)	(0.0055)	(0.0054)	(0.0055)	(0.0073)	(0.0069)	(0.0067)
R^2	0.011	0.030	0.033	0.035	0.053	0.128	0.143
	Panel C. Hourly Wage Rate						
Rent Stabilized	-0.082***	-0.061***	-0.053***	-0.060***	-0.086***	-0.076**	-0.071**
	(0.025)	(0.022)	(0.019)	(0.019)	(0.032)	(0.032)	(0.032)
R^2	0.138	0.199	0.268	0.272	0.292	0.315	0.332
Sub-borough and Year FE	✓	✓	✓	✓	✓	✓	✓
Demographic Control		✓	✓	✓	✓	✓	✓
Education Control			✓	✓	✓	✓	✓
Other Household Control				✓	✓	✓	✓
Housing and Neighborhood Control					✓	✓	✓
Industry Control						✓	✓
Occupation Control							✓

Notes: Standard errors are in parentheses and clustered at sub-borough level. * p<0.10, ** p<0.05, *** p<0.010. Demographic controls include gender, race and ethnicity, age, marital status, co-residence with parents, birth place. Other household controls include presence of children, presence of non-relatives, household size, and spouse's education and work status for married renters. Housing and neighborhood controls particularly include factors affecting the rent stabilization eligibility such as year a building is built and number of units in a building among others. Full results containing coefficients of other variables are available upon request.

Table 1.5: First Stage of Instrumental Variable Estimation

	IV1	IV2	IV3
	Panel A. Labor Force Participation		
Fraction of Vacant-for-Rent that are Stabilized	0.47***		0.33***
	(0.11)		(0.071)
Fraction of Vacant-for-Rent that are Market-Rate		-0.40***	-0.14*
		(0.11)	(0.074)
R^2	0.756	0.756	0.756
F Statistics	37.425	26.303	19.866
Over-Identifying Test			0.8339
	Panel B. Unemployment		
Fraction of Vacant-for-Rent that are Stabilized	0.49***		0.34***
	(0.11)		(0.070)
Fraction of Vacant-for-Rent that are Market-Rate		-0.42***	-0.15*
		(0.11)	(0.080)
R^2	0.750	0.750	0.750
F Statistics	35.962	25.395	19.238
Over-Identifying Test			0.3631
	Panel C. Hourly Wage Rate		
Fraction of Vacant-for-Rent that are Stabilized	0.47***		0.33***
	(0.11)		(0.076)
Fraction of Vacant-for-Rent that are Market-Rate		-0.40***	-0.14*
		(0.11)	(0.086)
R^2	0.748	0.748	0.748
F Statistics	34.191	24.934	18.013
Over-Identifying Test			0.9145

Notes: Standard errors are in parentheses and clustered at sub-borough level. * p<0.10, ** p<0.05, *** p<0.010. Number of observations (N) in each panel varies because the outcome variables are different. For example, comparing with panel A, the N is smaller in panel B because tenants who are not in the labor force are coded as not applicable for answering question to unemployed or not. All controls include sub-borough and year fixed effect, demographic, education, other household, housing and neighborhood, industry and occupation control variables included in the last column of Table 1.4. Full results containing coefficients of other variables are available upon request.

Table 1.6: Effect of Rent Stabilization on Labor Market Outcomes

(Instrumental Variable Estimation)

	Original OLS	IV1	IV2	IV3
	Panel A. Labor Force Participation			
Rent Stabilized	0.010	0.039	0.043	0.040
	(0.0073)	(0.037)	(0.040)	(0.036)
R^2	0.561	0.555	0.555	0.555
F Statistics		37.425	26.303	19.866
Over-Identifying Test				0.8339
	Panel B. Unemployment			
Rent Stabilized	0.023***	0.12***	0.14***	0.12***
	(0.0063)	(0.037)	(0.044)	(0.037)
R^2	0.132	0.091	0.073	0.088
F Statistics		35.962	25.395	19.238
Over-Identifying Test				0.3631
	Panel C. Hourly Wage Rate			
Rent Stabilized	-0.072**	-0.17	-0.16	-0.17
	(0.033)	(0.13)	(0.14)	(0.13)
R^2	0.322	0.312	0.312	0.312
F Statistics		34.191	24.934	18.013
Over-Identifying Test				0.9145

Notes: Standard errors are in parentheses and clustered at sub-borough level. * p<0.10, ** p<0.05, *** p<0.010. All controls include sub-borough and year fixed effect, demographic, education, other household, housing and neighborhood, industry and occupation control variables included in the last column of Table 1.4. Full results containing coefficients of other variables are available upon request.

Table 1.7: Rent Stabilization: Heterogenous Treatment Effects

	White		Skill		Age		Tenure	
	Yes	No	Low	High	26-40	41-55	Short	Long
				Panel A. Labor Force Participation				
Rent Stabilized	0.039	-0.0052	-0.14	0.076	0.029	-0.0013	-0.082	-0.12
	(0.041)	(0.090)	(0.12)	(0.047)	(0.064)	(0.057)	(0.21)	(0.11)
R^2	0.578	0.560	0.586	0.428	0.519	0.640	0.500	0.602
F Statistics	50.745	15.782	13.579	37.002	24.695	36.247	3.673	17.720
				Panel B. Unemployment				
Rent Stabilized	0.15***	0.035	0.088	0.14***	0.046	0.13**	-0.029	0.25
	(0.029)	(0.13)	(0.13)	(0.037)	(0.059)	(0.057)	(0.18)	(0.15)
R^2	0.031	0.152	0.132	0.051	0.131	0.114	0.128	0.004
F Statistics	49.359	13.886	12.634	33.091	23.608	38.393	2.931	13.879
				Panel C. Hourly Wage Rate				
Rent Stabilized	-0.15	-0.24	-0.29	-0.063	-0.18	-0.29	-0.13	-0.21
	(0.16)	(0.30)	(0.33)	(0.14)	(0.21)	(0.20)	(0.64)	(0.37)
R^2	0.227	0.273	0.204	0.217	0.314	0.314	0.337	0.301
F Statistics	43.101	16.575	14.667	27.840	22.149	34.838	3.463	12.066

Notes: Standard errors are in parentheses and clustered at sub-borough level. * p<0.10, ** p<0.05, *** p<0.010. Low skill tenants are those who do not have a college degree while high skill tenants hold at least a college degree. Short tenants are those who have lived in their housing units for less than 5 years while long tenants have lived for at least 5 years. All controls include sub-borough and year fixed effect, demographic, education, other household, housing and neighborhood, industry and occupation control variables included in the last column of Table 1.4. Full results containing coefficients of other variables are available upon request.

Table 1.8: Legal Status versus Self-Reported Regulation Status

Self-Reported Status	Legal Status		
	Market-rate	Rent-stabilized	Total
Rent-controlled	70	474	544
	1.79%	8.60%	5.77%
Rent-stabilized	136	1409	1545
	3.48%	25.55%	16.40%
Market-rate	2317	1338	3655
	59.29%	24.27%	38.79%
Don't Know	621	1198	1819
	15.89%	21.73%	19.31%
Not Reported	764	1095	1859
	19.55%	19.86%	19.73%
Total	3908	5514	9422
	100.00%	100.00%	100.00%

Notes: Data comes from pooled 2002 and 2005 waves of the New York City Housing and Vacancy Survey (NYCHVS).

Table 1.9: Effect of Rent Stabilization Awareness on Unemployment (Full Sample)

	OLS	Probit	Logit
Aware	0.019**	0.023**	0.022**
	(0.0072)	(0.0097)	(0.0088)
Misinformed	0.015	0.017	0.015
	(0.0094)	(0.011)	(0.010)
R^2	0.130	0.104	0.103
All Other Controls	✓	✓	✓

Notes: Standard errors are in parentheses and clustered at sub-borough level. * p<0.10, ** p<0.05, *** p<0.010. A rent-stabilized tenant is *aware* if the self-reported regulation status is either rent-controlled or rent-stabilized. A rent-stabilized tenant is *misinformed* if the self-reported regulation status is market-rate. Unemployment is a binary variable, i.e. Y=1 if a renter is currently unemployed. All controls include sub-borough and year fixed effect, demographic, education, other household, housing and neighborhood, industry and occupation control variables included in the last column of Table 1.4. Full results containing coefficients of other variables are available upon request.

Table 1.10: First Stage of Instrumental Variable for Policy Awareness

	IV (1)		IV (2)	
	Aware	Misinformed	Aware	Misinformed
Fraction of Vacant-for-Rent that are Stabilized	0.80***	0.045	0.80***	0.050
	(0.12)	(0.055)	(0.12)	(0.056)
Local Rent Regulation Knowledge: Aware	0.28**	-0.14	0.54***	-0.49***
	(0.12)	(0.10)	(0.13)	(0.11)
Local Rent Regulation Knowledge: Misinformed	-0.33***	0.43***		
	(0.086)	(0.091)		
Local Rent Regulation Knowledge: Don't Know			0.31***	-0.42***
			(0.11)	(0.11)
R^2	0.242	0.148	0.241	0.146

Notes: Standard errors are in parentheses and clustered at sub-borough level. * p<0.10, ** p<0.05, *** p<0.010. All controls include sub-borough and year fixed effect, demographic, education, other household, housing and neighborhood, industry and occupation control variables included in the last column of Table 1.4. Full results containing coefficients of other variables are available upon request.

Table 1.11: Effect of Policy Awareness on Unemployment

(Instrumental Variable Estimation)

	Original OLS	IV (1)	IV (2)
Aware	0.019**	0.11***	0.11***
	(0.0072)	(0.043)	(0.043)
Misinformed	0.015	0.046	0.025
	(0.0094)	(0.095)	(0.10)
$\overline{Y}$	0.061	0.061	0.061
N	6157	6157	6157
R^2	0.131	0.107	0.105
F Statistics		10.764	8.733

Notes: Standard errors are in parentheses and clustered at sub-borough level. * p<0.10, ** p<0.05, *** p<0.010. A rent-stabilized tenant is *aware* if the self-reported regulation status is either rent-controlled or rent-stabilized. A rent-stabilized tenant is *misinformed* if the self-reported regulation status is market-rate. Unemployment is a binary variable, i.e. Y=1 if a renter is currently unemployed. All controls include sub-borough and year fixed effect, demographic, education, other household, housing and neighborhood, industry and occupation control variables included in the last column of Table 1.4. Full results containing coefficients of other variables are available upon request.

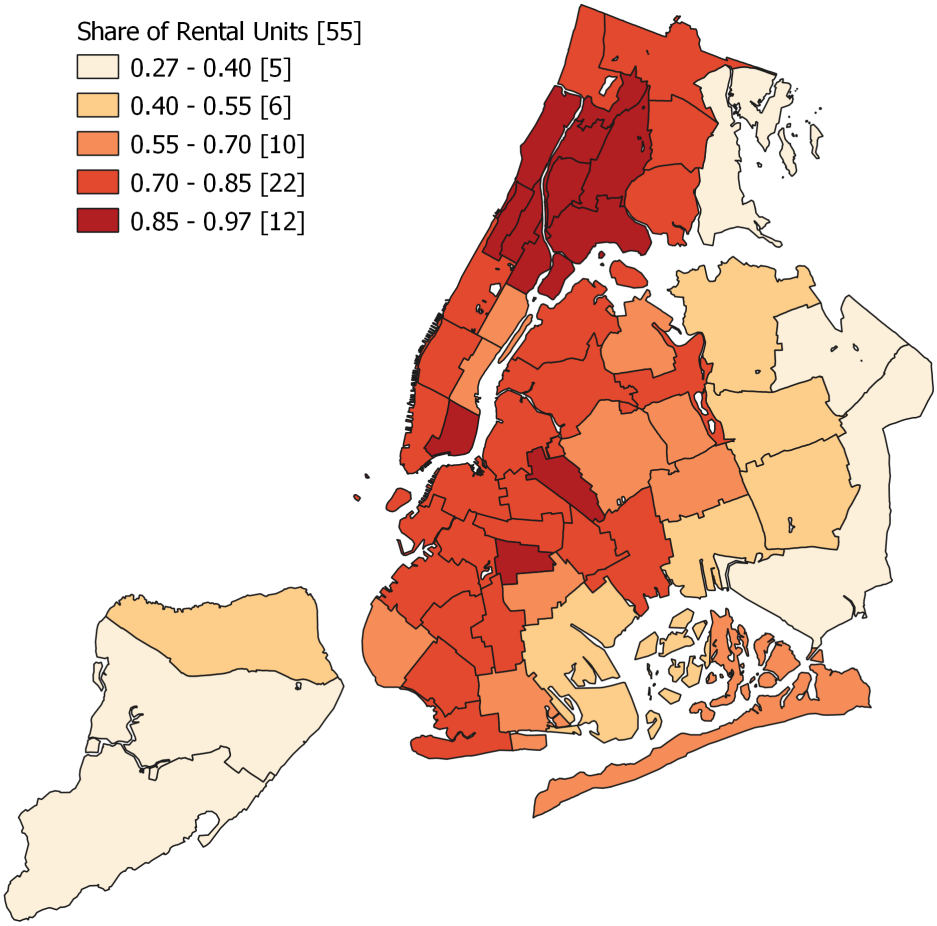
Table 1.12: Rent Discount and Mobility as Suggestive Mechanisms

	Rent Stab. Dummy	Rent Discount	+ Mobility	Alter. Mobility
Rent Stabilized	0.022***			
	(0.0070)			
Rent Discount		0.019***	0.016**	0.017***
		(0.0060)	(0.0061)	(0.0063)
Tenure of Residence			0.00061	
			(0.00054)	
Tenure: 2-4 Years				-0.0097
				(0.0071)
Tenure: 5-7 Years				0.0029
				(0.0086)
Tenure: 8-10 Years				-0.00085
				(0.011)
Tenure: 10+ Years				0.0041
				(0.012)
R^2	0.149	0.135	0.135	0.136

Notes: Standard errors are in parentheses and clustered at sub-borough level. * p<0.10, ** p<0.05, *** p<0.010. Full results containing coefficients of other variables are available upon request. Rent discount is obtained from Jiang et al. (2019).

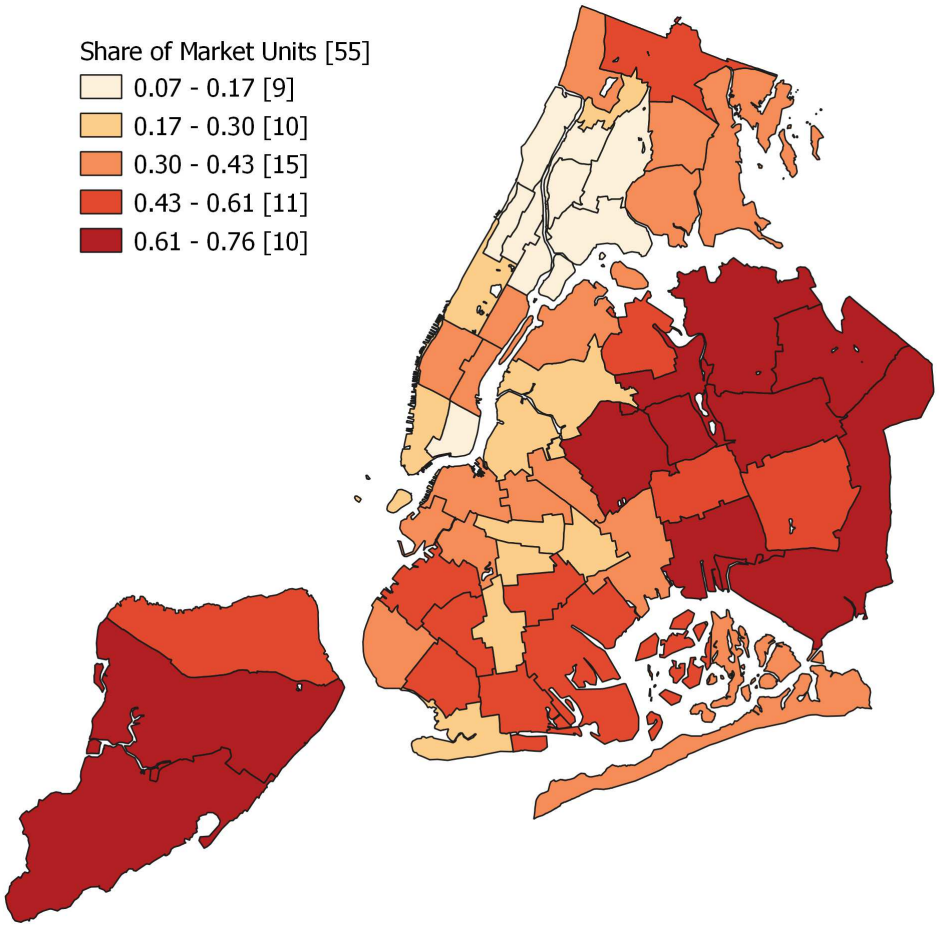
1.11 Appendix

Figure 1.10: Spatial Concentration of Rental Occupied Units in NYC



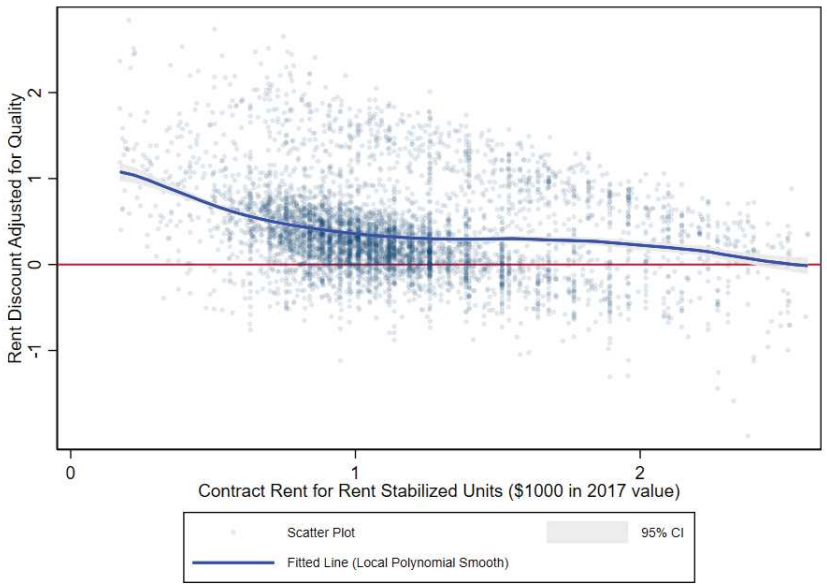
Notes: Author's calculation based on pooled New York City Housing and Vacancy Survey (NYCHVS) 2002 and 2005.

Figure 1.11: Spatial Concentration of Private Market-Rate Units in NYC



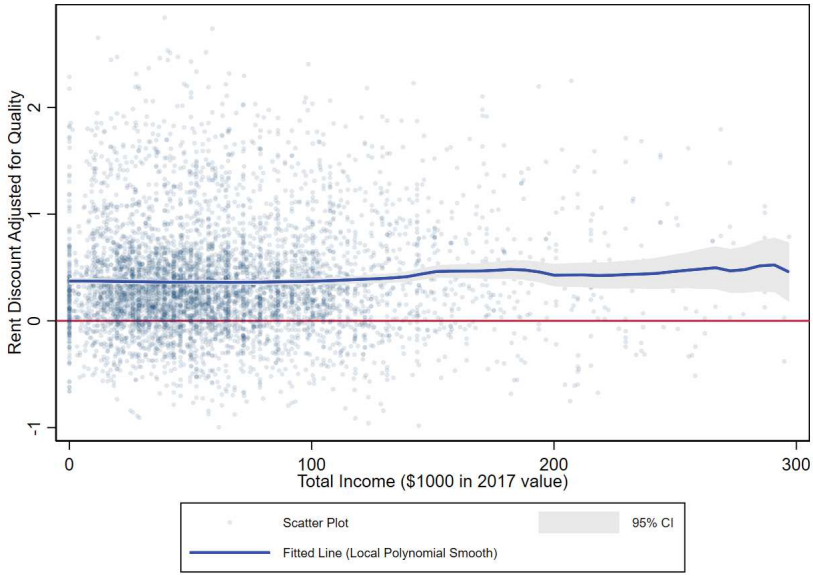
Notes: Author's calculation based on pooled New York City Housing and Vacancy Survey (NYCHVS) 2002 and 2005.

Figure 1.12: Rent Discount v.s. Contract Rent



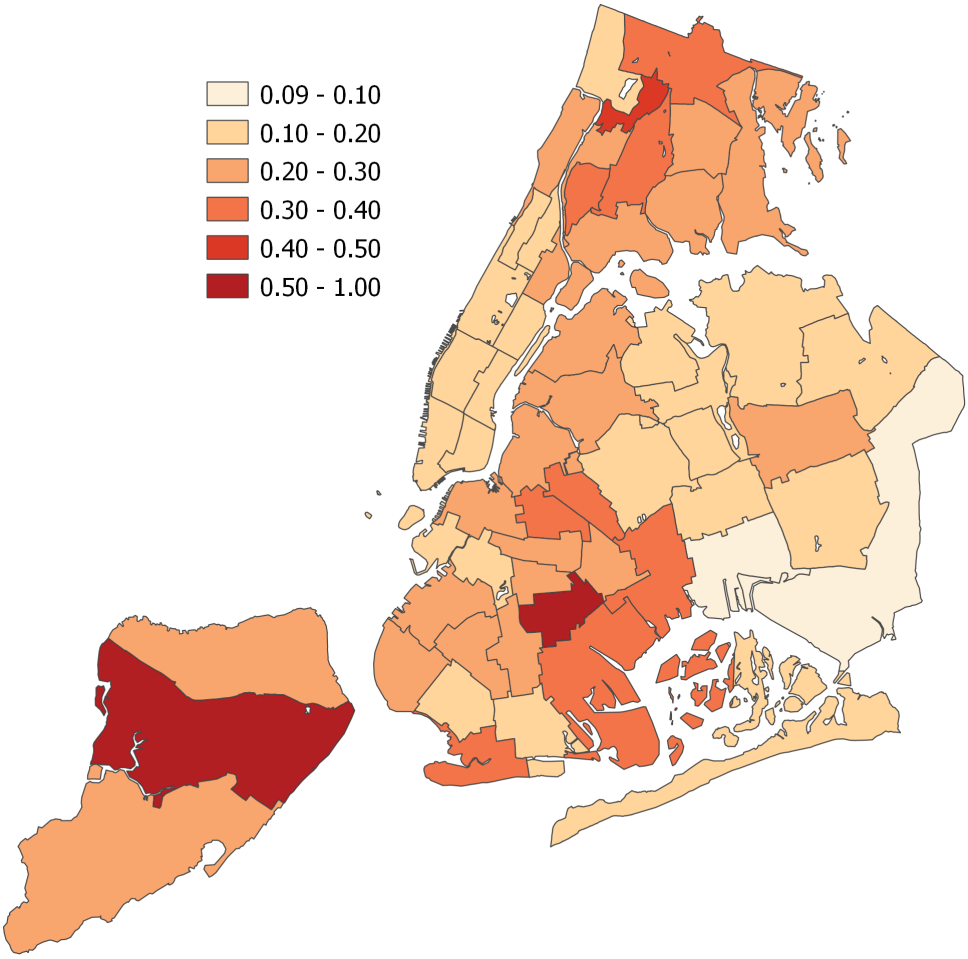
Notes: Calculation is based on the rent discount estimated in Jiang et al. (2019) where different empirical hedonic pricing models are analyzed and compared. The unit of both Y-axis and X-axis are in $1000 2017 value.

Figure 1.13: Rent Discount v.s. Total Household Income



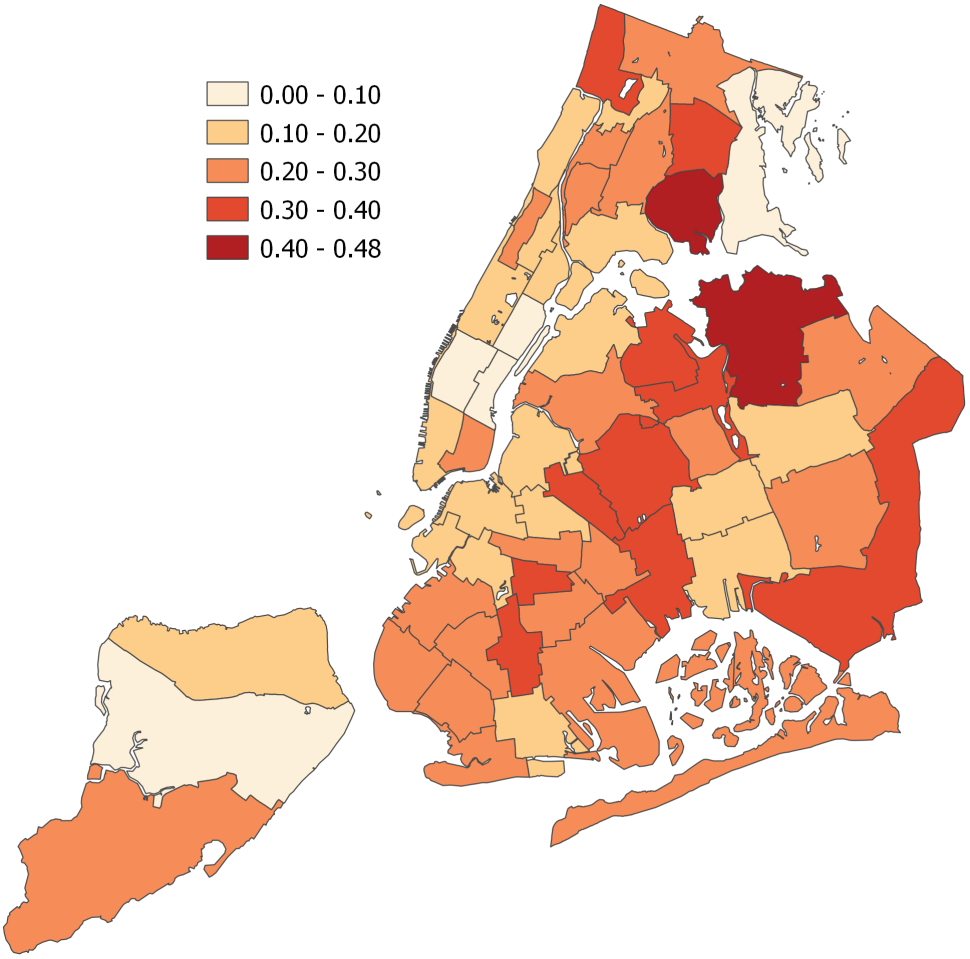
Notes: Calculation is based on the rent discount estimated in Jiang et al. (2019) where different empirical hedonic pricing models are analyzed and compared. The unit of both Y-axis and X-axis are in $1000 2017 value.

Figure 1.14: Spatial Concentration of Policy Awareness (Misinform)



Notes: Author's calculation based on pooled New York City Housing and Vacancy Survey (NYCHVS) 2002 and 2005.

Figure 1.15: Spatial Concentration of Policy Awareness (Don't Know)



Notes: Author's calculation based on pooled New York City Housing and Vacancy Survey (NYCHVS) 2002 and 2005.

Table 1.13: Summary Statistics of Household by Rent Stabilization

	Rent Stabilized				Market Rental			
	mean	sd	min	max	mean	sd	min	max
Currently In the Labor Force	0.889	0.31	0.00	1.00	0.888	0.32	0.00	1.00
Currently Unemployed	0.063	0.24	0.00	1.00	0.052	0.22	0.00	1.00
HHH Wage Rate	29.017	25.84	0.45	172.18	30.697	27.20	0.45	172.18
High School Dropout	0.166	0.37	0.00	1.00	0.154	0.36	0.00	1.00
High School Graduate	0.233	0.42	0.00	1.00	0.241	0.43	0.00	1.00
Some College No Degree	0.199	0.40	0.00	1.00	0.206	0.40	0.00	1.00
College Degree and Above	0.402	0.49	0.00	1.00	0.399	0.49	0.00	1.00
Female	0.511	0.50	0.00	1.00	0.449	0.50	0.00	1.00
Age	38.925	7.80	26.00	54.00	38.194	7.77	26.00	54.00
White, Non-Hispanic	0.356	0.48	0.00	1.00	0.414	0.49	0.00	1.00
Black, Non-Hispanic	0.220	0.41	0.00	1.00	0.225	0.42	0.00	1.00
Hispanic	0.313	0.46	0.00	1.00	0.213	0.41	0.00	1.00
Asian	0.103	0.30	0.00	1.00	0.140	0.35	0.00	1.00
Married	0.349	0.48	0.00	1.00	0.438	0.50	0.00	1.00
Born in NYC	0.215	0.41	0.00	1.00	0.243	0.43	0.00	1.00
Born in U.S. (Non-NYC)	0.141	0.35	0.00	1.00	0.132	0.34	0.00	1.00
Born in Other Country	0.467	0.50	0.00	1.00	0.446	0.50	0.00	1.00
Birth Place Missing	0.177	0.38	0.00	1.00	0.179	0.38	0.00	1.00
Any Young Child in HH	0.160	0.37	0.00	1.00	0.195	0.40	0.00	1.00
Any Child in HH	0.364	0.48	0.00	1.00	0.408	0.49	0.00	1.00
Co-reside with Parents	0.037	0.19	0.00	1.00	0.035	0.18	0.00	1.00
HH Size	2.397	1.43	1.00	11.00	2.688	1.58	1.00	15.00
Any Non-relative in HH	0.108	0.31	0.00	1.00	0.132	0.34	0.00	1.00
HHH Salary Income, cond. work	52.563	40.85	0.09	215.23	56.978	44.38	0.37	215.23
HHH Non-Labor Income	2.948	10.28	0.00	68.73	3.119	10.59	0.00	68.73
Other HH Mem. Tot. Inc.	20.037	31.36	-0.00	170.46	26.745	37.06	-0.00	170.46

Table 1.14: Summary Statistics of Housing Characteristics

	Rent Stabilized				Market Rental			
	mean	sd	min	max	mean	sd	min	max
Monthly Contract Rent	1.176	0.50	0.02	3.67	1.489	0.75	0.17	3.67
Monthly Gross Rent	1.271	0.50	0.14	3.81	1.618	0.77	0.17	3.81
Tenure of Residence	6.998	5.89	1.00	27.00	5.048	4.62	1.00	27.00
Num. of Rooms	3.174	1.16	1.00	8.00	3.908	1.30	1.00	8.00
Num. of Bedrooms	1.341	0.81	0.00	7.00	1.834	0.92	0.00	8.00
Num. of Unit Problems	0.909	1.22	0.00	6.00	0.463	0.87	0.00	6.00
Num. of Building Problems	0.257	0.76	0.00	5.00	0.184	0.67	0.00	5.00

Table 1.15: Effect of Rent Stabilization on Labor Force Participation and Unemployment (Alternative Model Specifications, Full Sample)

	LPM	Probit	Logit
	Panel A. Labor Force Participation		
Rent Stabilized	0.0099	0.0053	0.0078
	(0.0072)	(0.012)	(0.015)
R^2	0.565	0.176	0.178
	Panel B. Unemployment		
Rent Stabilized	0.023***	0.036***	0.034***
	(0.0067)	(0.011)	(0.0099)
R^2	0.143	0.114	0.115

Notes: Standard errors are in parentheses and clustered at sub-borough level. * p<0.10, ** p<0.05, *** p<0.010. Labor force participation and unemployment are binary variables, i.e. Y=1 if a renter participates in currently the labor force in Panel A and Y=1 if a renter is currently unemployed in Panel B. Demographic controls include gender, race and ethnicity, age, marital status, co-residence with parents, birth place. Other household controls include presence of children, presence of non-relatives, household size, and spouse's education and work status for married renters. Housing and neighborhood controls particularly include factors affecting the rent stabilization eligibility such as year a building is built and number of units in a building among others. Full results containing coefficients of other variables are available upon request.

Table 1.16: Summary Statistics of Household by Policy Awareness

	Aware		Misinformed		Market-Rate	
	mean	sd	mean	sd	mean	sd
Female	0.525	0.50	0.496	0.50	0.449	0.50
Age	39.445	7.97	38.142	7.81	38.194	7.77
White, Non-Hispanic	0.474	0.50	0.255	0.44	0.414	0.49
Black, Non-Hispanic	0.189	0.39	0.233	0.42	0.225	0.42
Hispanic	0.249	0.43	0.383	0.49	0.213	0.41
Asian	0.076	0.26	0.123	0.33	0.140	0.35
Married	0.319	0.47	0.401	0.49	0.438	0.50
Born in NYC	0.300	0.46	0.257	0.44	0.243	0.43
Born in U.S. (Non-NYC)	0.245	0.43	0.140	0.35	0.132	0.34
Born in Other Country	0.431	0.50	0.586	0.49	0.446	0.50
Birth Place Missing	0.024	0.15	0.017	0.13	0.179	0.38
Any Young Child in HH	0.129	0.34	0.221	0.42	0.195	0.40
Any Child in HH	0.311	0.46	0.446	0.50	0.408	0.49
Co-reside with Parents	0.031	0.17	0.047	0.21	0.035	0.18
HH Size	2.221	1.34	2.675	1.52	2.688	1.58
Any Non-relative in HH	0.111	0.31	0.119	0.32	0.132	0.34
High School Dropout	0.114	0.32	0.173	0.38	0.154	0.36
High School Graduate	0.165	0.37	0.264	0.44	0.241	0.43
Some College No Degree	0.202	0.40	0.220	0.41	0.206	0.40
College Degree and Above	0.519	0.50	0.344	0.48	0.399	0.49
Currently In the Labor Force	0.914	0.28	0.891	0.31	0.888	0.32
Currently Unemployed	0.071	0.26	0.065	0.25	0.052	0.22
Annual Hours of Work, cond. work	1968.118	677.59	1970.438	653.40	2011.216	664.04
HHH Salary Income, cond. work	59.590	43.87	49.784	38.43	56.978	44.38
HHH Wage Rate	32.630	26.94	27.489	24.47	30.697	27.20
HHH Non-Labor Income	3.677	11.77	2.608	9.38	3.119	10.59
Other HH Mem. Tot. Inc.	21.214	33.34	23.064	32.73	26.745	37.06

Table 1.17: Summary Statistics of Housing by Policy Awareness

	Aware		Misinformed		Market-Rate	
	mean	sd	mean	sd	mean	sd
Monthly Contract Rent	1.208	0.51	1.200	0.53	1.489	0.75
Monthly Gross Rent	1.297	0.51	1.299	0.54	1.618	0.77
Tenure of Residence	8.153	6.35	6.641	5.64	5.048	4.62
Num. of Rooms	3.070	1.17	3.282	1.19	3.908	1.30
Num. of Bedrooms	1.270	0.79	1.430	0.84	1.834	0.92
Nbhd Rate: Excellent	0.165	0.37	0.117	0.32	0.175	0.38
Nbhd Rate: Good	0.509	0.50	0.514	0.50	0.489	0.50
Nbhd Rate: Fair	0.254	0.44	0.301	0.46	0.138	0.34
Nbhd Rate: Poor	0.052	0.22	0.055	0.23	0.019	0.14
Nbhd Rate: Missing	0.020	0.14	0.013	0.12	0.180	0.38

Table 1.18: Summary Statistics of Household: Aware v.s. Private Market

	Aware	N	Market-Rate	N	Difference
Female	0.525	1883	0.449	3908	0.075***
Age	39.445	1883	38.194	3908	1.250***
White, Non-Hispanic	0.474	1883	0.414	3908	0.060***
Black, Non-Hispanic	0.189	1883	0.225	3908	-0.036***
Hispanic	0.249	1883	0.213	3908	0.036***
Asian	0.076	1883	0.140	3908	-0.065***
Married	0.319	1883	0.438	3908	-0.118***
Born in NYC	0.300	1883	0.243	3908	0.057***
Born in U.S. (Non-NYC)	0.245	1883	0.132	3908	0.113***
Born in Other Country	0.431	1883	0.446	3908	-0.016
Birth Place Missing	0.024	1883	0.179	3908	-0.154***
Any Young Child in HH	0.129	1883	0.195	3908	-0.066***
Any Child in HH	0.311	1883	0.408	3908	-0.097***
Co-reside with Parents	0.031	1883	0.035	3908	-0.004
HH Size	2.221	1883	2.688	3908	-0.467***
Any Non-relative in HH	0.111	1883	0.132	3908	-0.021**
High School Dropout	0.114	1883	0.154	3908	-0.040***
High School Graduate	0.165	1883	0.241	3908	-0.076***
Some College No Degree	0.202	1883	0.206	3908	-0.004
College Degree and Above	0.519	1883	0.399	3908	0.120***
Currently In the Labor Force	0.914	1883	0.888	3908	0.026***
Currently Unemployed	0.071	1722	0.052	3471	0.019***
Annual Hours of Work, cond. work	1968.118	1727	2011.216	3476	-43.099**
HHH Salary Income, cond. work	59.590	1655	56.978	3333	2.613**
HHH Wage Rate	32.630	1655	30.697	3333	1.933**
HHH Non-Labor Income	3.677	1883	3.119	3908	0.558*
Other HH Mem. Tot. Inc.	21.214	1883	26.745	3908	-5.531***

Table 1.19: Summary Statistics of Household: Misinformed v.s. Private Market

	Misinformed	N	Market-Rate	N	Difference
Female	0.496	1338	0.449	3908	0.046***
Age	38.142	1338	38.194	3908	-0.052
White, Non-Hispanic	0.255	1338	0.414	3908	-0.159***
Black, Non-Hispanic	0.233	1338	0.225	3908	0.008
Hispanic	0.383	1338	0.213	3908	0.169***
Asian	0.123	1338	0.140	3908	-0.017
Married	0.401	1338	0.438	3908	-0.037**
Born in NYC	0.257	1338	0.243	3908	0.015
Born in U.S. (Non-NYC)	0.140	1338	0.132	3908	0.007
Born in Other Country	0.586	1338	0.446	3908	0.140***
Birth Place Missing	0.017	1338	0.179	3908	-0.162***
Any Young Child in HH	0.221	1338	0.195	3908	0.026**
Any Child in HH	0.446	1338	0.408	3908	0.038**
Co-reside with Parents	0.047	1338	0.035	3908	0.013**
HH Size	2.675	1338	2.688	3908	-0.013
Any Non-relative in HH	0.119	1338	0.132	3908	-0.013
High School Dropout	0.173	1338	0.154	3908	0.018
High School Graduate	0.264	1338	0.241	3908	0.023*
Some College No Degree	0.220	1338	0.206	3908	0.013
College Degree and Above	0.344	1338	0.399	3908	-0.055***
Currently In the Labor Force	0.891	1338	0.888	3908	0.003
Currently Unemployed	0.065	1192	0.052	3471	0.013*
Annual Hours of Work, cond. work	1970.438	1197	2011.216	3476	-40.778*
HHH Salary Income, cond. work	49.784	1159	56.978	3333	-7.194***
HHH Wage Rate	27.489	1159	30.697	3333	-3.208***
HHH Non-Labor Income	2.608	1338	3.119	3908	-0.511
Other HH Mem. Tot. Inc.	23.064	1338	26.745	3908	-3.681***

Table 1.20: First Stage of Alternative Instrumental Variable (Unemployment)

	IV4	IV5	IV6
Total Number of Vacant-for-Rent Units (1000)	-0.014***	0.011***	
	(0.0035)	(0.0025)	
Total Number of Vacant-for-Rent and Stabilized Units (1000)	0.028***		0.013***
	(0.0060)		(0.0027)
Total Number of Vacant-for-Rent and Market-Rate Units (1000)		-0.028***	-0.017***
		(0.0065)	(0.0042)
$\overline{Y}$	0.585	0.585	0.585
N	8335	8335	8335
R^2	0.751	0.751	0.752
All Other Controls	✓	✓	✓

Notes: Standard errors are in parentheses and clustered at sub-borough level. * p<0.10, ** p<0.05, *** p<0.010. Full results containing coefficients of other variables are available upon request.

Table 1.21: Effect of Rent Stabilization on Unemployment (Alternative IV)

	Original Share IV			Alternative IV		
	IV1	IV2	IV3	IV4	IV5	IV6
Stabilized	0.12***	0.14***	0.12***	0.099**	0.11***	0.11***
	(0.039)	(0.045)	(0.039)	(0.040)	(0.042)	(0.040)
$\overline{Y}$	0.058	0.058	0.058	0.058	0.058	0.058
N	8335	8335	8335	8335	8335	8335
R^2	0.092	0.073	0.088	0.104	0.094	0.099
F Statistics	34.364	24.401	18.330	19.192	16.719	19.130
Over-Identifying Test			0.3631	0.2579	0.2306	0.1842
All Other Controls	✓	✓	✓	✓	✓	✓

Notes: Standard errors are in parentheses and clustered at sub-borough level. * p<0.10, ** p<0.05, *** p<0.010. Full results containing coefficients of other variables are available upon request.

Table 1.22: Effect of Policy Awareness on Unemployment

(Including Missing Values for Policy Awareness)

	Original OLS	IV
Aware	0.021**	0.12***
	(0.0071)	(0.040)
Misinformed	0.016	0.079
	(0.0092)	(0.112)
Missing	0.014	0.203
	(0.0087)	(0.245)
$\overline{Y}$	0.061	0.061
N	7117	7117
R^2	0.128	0.126
F Statistics		4.13

Notes: Standard errors are in parentheses and clustered at sub-borough level. * p<0.10, ** p<0.05, *** p<0.010. A rent-stabilized tenant is *aware* if the self-reported regulation status is either rent-controlled or rent-stabilized. A rent-stabilized tenant is *misinformed* if the self-reported regulation status is market-rate. Unemployment is a binary variable, i.e. Y=1 if a renter is currently unemployed. All controls include sub-borough and year fixed effect, demographic, education, other household, housing and neighborhood, industry and occupation control variables included in the last column of Table 1.4. Full results containing coefficients of other variables are available upon request.

Supplementary Info about Institution

Table 1.23: Rent Stabilization v.s. Rent Control in NYC

	Rent Stabilization	Rent Control
Enact Time	1969	1943
Current Stock (2017)	966,000 out of 2,183,064 (44%)	22,000 of 2,183,064 (1%)
Rent Increase Operation	Rent Guidelines Board (RGB)	Maximum Base Rent (MBR)
Succession Right	Yes	Yes
Criterion 1: Construction Time	(1) Between 2/1/1947 and 1/1/1974 (2) Post-1974 units with tax benefits	Only before 2/1/1947
Criterion 2: Number of Units	Mainly ≥ 6	Mainly ≥ 3
Criterion 3: Move-in Time	Move-in after 6/30/1971 for units built before 2/1/1947	Move-in before 7/1/1971
Vacancy De-regulation	(1) Deregulation if *exceeds* High-Rent threshold (2) "Vacancy Bonus" if *below* High-Rent threshold	(1) Deregulation if < 6 units (2) Become rent stabilized if ≥ 6 units
Other De-regulation	(1) Post Condo or Co-op conversion (2) Exceeds High-Income threshold (3) Tax benefit expires	

Notes: Author's own summarization based on legal regulation documents.

Table 1.24: Permitted Annual Increase of Rent Stabilized Apartments (1978-2020)

Renewal Leases Time	One Year Lease	Two Year Lease
10/1/19 to 9/30/20	1.5%	2.5%
10/1/18 to 9/30/19	1.5%	2.5%
10/1/17 to 9/30/18	1.25%	2%
10/1/16 to 9/30/17	0%	2%
10/1/15 to 9/30/16	0%	2%
10/1/14 to 9/30/15	1%	2.75%
10/1/13 to 9/30/14	4%	7.75%
10/1/12 to 9/30/13	2%	4%
10/1/11 to 9/30/12	3.75%	7.25%
10/1/10 to 9/30/11	2.25%	4.5%
10/1/09 to 9/30/10	3%	6%
10/1/08 to 9/30/09	4.5%	8.5%
10/1/07 to 9/30/08	3%	5.75%
10/1/06 to 9/30/07	4.25%	7.25%
10/1/05 to 9/30/06	2.75%	5.5%
10/1/04 to 9/30/05	3.5%	6.5%
10/1/03 to 9/30/04	4.5%	7.5%
10/1/02 to 9/30/03	2%	4%
10/1/01 to 9/30/02	4%	6%
10/1/00 to 9/30/01	4%	6%
10/1/99 to 9/30/00	2%	4%
10/1/98 to 9/30/99	2%	4%
10/1/97 to 9/30/98	2%	4%
10/1/96 to 9/30/97	5%	7%
10/1/95 to 9/30/96	2%	4%
10/1/94 to 9/30/95	2%	4%
10/1/93 to 9/30/94	3%	5%
10/1/92 to 9/30/93	3%	5%
10/1/91 to 9/30/92	4%	6.5%
10/1/90 to 9/30/91	4.5%	7%
10/1/89 to 9/30/90	5.5%	9%
10/1/88 to 9/30/89	6%	9%
10/1/87 to 9/30/88	3%	6.5%
10/1/86 to 9/30/87	6%	9%
10/1/85 to 9/30/86	4%	6.5%
10/1/84 to 9/30/85	6%	9%
10/1/83 to 9/30/84	4%	7%
10/1/82 to 9/30/83	4%	7%
10/1/81 to 9/30/82	10%	13%
7/1/80 to 6/30/81	11%	14%
7/1/79 to 6/30/80	8.5%	12%
7/1/78 to 6/30/79	2.5%	2%

Notes: information is obtained from https://www1.nyc.gov/assets/rentguidelinesboard/pdf/guidelines/aptorders.pdf.

Chapter 2

Earned Income Tax Credit, Maternal Behavioral Response and Child Development

[1]

[1]I gratefully acknowledge my advisers: Robert Moffitt, Yingyao Hu, Luis Quintero, and Richard Spady for their encouragement, guidance and support. For helpful comments and conversations, I also thank from: Burt Barnow, Eric French, Chris Herbst, Maggie Jones, Sandra Newman, Erdal Tekin, James Ziliak along with seminar participants at 2018 PSID Annual User Conference, 2018 APPAM Annual Fall Research Conference, 2018 NTA Annual Conference on Taxation, 2018 SEA Annual Conference, 2019 SOLE Annual Meeting. All errors are my own.

2.1 Abstract

This paper investigates the impact of the Earned Income Tax Credit (EITC) on children's cognitive and non-cognitive skills. Given that the theoretical impact of the EITC on child development is ambiguous, this paper particularly emphasizes the underlying mechanisms. Using data from the Panel Study of Income Dynamics (PSID) and its Child Development Supplement (CDS) in combination with an augmented sample of all non-parental child care histories, increased EITC generosity is found to negatively affect cognitive skills of children from single-mother families. No significant effect on children from married-mother families is found. Positive effects are found regarding children's non-cognitive skills, irrespective of family structure. Furthermore, several mechanisms are shown to be important. First, there is strong evidence that both single-mother and married-mother families increase their non-parental child care usage when facing higher EITC generosity, but this is enacted in different forms. Single mothers mainly utilize relatives for child care, while married mothers tend to utilize formal child care programs. Second, single mothers tend to shift from actively engaged time with children to passively accessible time. Third, single mothers are found to have increased psychological distress. Both the time input and psychological distress channels are not salient among married-mothers. Finally, the analysis shows that increased goods inputs, measured by income, may not directly correspond with increased child-specific expenditures.

2.2 Introduction

The Earned Income Tax Credit (EITC) is one of the most important welfare programs in the U.S., as it has proven to contribute to increases in employment, labor supply, and earning, as well as decreases in welfare use, and poverty among families.[2] Given that primary recipients of EITC are working parents with children, a small but growing group of studies have explored the potential intergenerational effects of the EITC on infant health (Hoynes et al., 2015; Strully et al., 2010), children's academic achievements (Dahl and Lochner, 2012), and their education and labor market outcomes in the long run (Bastian and Michelmore, 2018; Maxfield, 2013). However, few studies have investigated the impact of the EITC on children's cognitive and non-cognitive skill formation. Moreover, there have been even fewer studies that discuss underlying mechanisms given that the theoretical impact of the EITC on child outcome is ambiguous in sign.[3]

It has been well-documented that human capital skills developed in early stages, both cognitive and non-cognitive, have long-term positive impacts on socioeconomic outcomes in the life-cycle.[4] Among all of the inputs that shape skills, maternal time

[2]See Hotz and Scholz (2003), Grogger (2003) and **?** for recent reviews.

[3]Unlike unconditional cash assistance programs, the EITC has a strong "pro-work" nature, leading to its bipartisan support from the policy side. However, an EITC-induced labor supply change does not necessarily imply an one-to-one mapping to the change of maternal time input into a child, which is a result of maternal utility maximization. Moreover, a mother may also switch to alternative child care arrangements when a child is young, the quality of which is often questionable. Therefore, from an *ex-ante* perspective, it is ambiguous to predict the overall effect of the EITC on child development. Meanwhile, an *ex-post* insignificant estimate of the effect of EITC on any child outcome might not be very meaningful either, if different mechanisms are simply acting in different directions and therefore offsetting each other.

[4]See Currie and Almond (2011); Heckman and Mosso (2014) for recent summaries.

investment on children has particularly positive effects and the rate of return is much larger in early stages as opposed to later stages (Cunha and Heckman, 2008; Cunha et al., 2010). In contrast, maternal employment and low-quality child care have been shown to be detrimental to children's cognitive development (such as Bernal (2008); Bernal and Keane (2011) among others).[5] Hence, an important question remains unclear is to understand the different mechanisms through which the EITC affects child development.

This paper uses the Panel Study of Income Dynamics (PSID) and its Child Development Supplement (CDS) in a novel way to address these issues. The PSID is suitable not only for its detailed longitudinal information about children's developmental outcomes but also for its unique feature called a *time diary*. This allows for a key differentiation between a mother's time with her child and her leisure time (Del Boca et al., 2017, 2014).[6] Moreover, to better understand the child care channel, I leverage the CDS data in an innovative way by transforming the retrospective cross-sectional information on non-parental child care arrangement to a complete panel history up until kindergarten for each sample child.

The empirical identification relies on a key measured - EITC generosity, which is defined as the maximum potential federal and state credit a child's family could

[5]None of these studies have been put in the context of the EITC.

[6]Most studies have been using maternal employment as a proxy measure for maternal time input in understanding child development (including Bernal (2008); Bernal and Keane (2011); Liu et al. (2010); Todd and Wolpin (2007)), which assume a monotonic one-to-one mapping between these two factors but ignore the possibility that a mother can sacrifice her leisure, maintain the same (or even increase) time with her child, and increase her labor supply. Such limitation is mainly due to the fact that most socioeconomic surveys lack precise measure of a mother's time allocation.

receive given their residence, family size, and tax year. EITC generosity does not depend on family income or parental marital status.[7] In recent studies (including Bastian and Michelmore (2018); Hoynes et al. (2015); Maxfield (2013)), this measure has been used to take advantage of a series of federal and state EITC changes as opposed to relying on one single reform.[8] The identification comes from three sources of exogenous variation: (1) the year the individual was born, which affects the generosity of the federal credit in a specific year; (2) the state of residence in each year, which leads to different state EITC generosity; (3) the number of children in the household in each year, which leads to larger benefits for larger households. Furthermore, alternative evaluation methods have been used including state fixed effect, individual fixed effect and value-added models. The results are robust and highlight differences on the basis of family structure.

The estimated results suggest that, on the one hand, EITC has negative impacts on cognitive skills for children from single-mother families as compared to married-mother families. On the other hand, EITC is also found to improve children's non-

[7]The actual eligibility and the amount of EITC are endogenous, because they are determined by household earning. Notably that the PSID does not have the actual amount of EITC an eligible family receives. However, all the relevant household demographic and income information allows me to simulate both the eligibility and the federal and state credit using NBER's TAXSIM.

[8]The majority of literature on the effect of the EITC on certain outcomes has been predominantly using quasi-experimental approach, particularly the Difference-in-Differences (DID) strategy, by exploiting the three major federal level expansion of EITC. The *Tax Reform Act of 1986 (TRA86)* which mainly increased the credit rate, has been most notably used by Eissa and Liebman (1996). The *1990 Omnibus Budget Reconciliation Act (OBRA90)* created the separate benefit schedules for 1 and 2 and more-child families, which has been used Boyd-Swan et al. (2016). The *1993 Omnibus Budget Reconciliation Act (OBRA93)*, which generated significant differences in the EITC benefits between families with *2 and more* children and those with *only 1* child, has been used most widely in the literature as it is regarded as the largest expansion to-date of the EITC by: Cowan and Tefft (2012) and Averett and Wang (2013) (women's smoking), Evans and Garthwaite (2014) (maternal mental and psychical health), and Hoynes et al. (2015) (infant health).

cognitive skills. Such a positive effect is salient for children from both single-mother families and married-mother families. The results are robust based on different empirical model specifications.

There are at least three channels that may help explain the effects of EITC on child outcomes. The first important channel is non-parental child care. Both single mothers and married mothers tend to increase child care usage at both extensive and intensive margins. However, single mothers are more likely to choose informal child care arrangements with relatives, which is usually thought to be of low quality and as potentially causing negative consequence on child development. In contrast, married mothers are more likely to use formal child care. Second, single mothers are more likely to change their time inputs by switching between time types. In other words, they reduce their actively engaged time but increase their accessible time. Such a shift has not been found among married mothers. Reduced engaged time may also be detrimental to children's cognitive development. Third, single mothers are in a worse position and experience more distress when faced with increased EITC generosity. However, the EITC has a modest positive effect on the psychological distress of married mothers.[9] The effects based on these three channels may help explain why children from single-mother families are in a worse position regarding their cognitive skills. Finally, it remains questionable whether increased family incomes directly relate to increased child-specific expenditure.

[9]See Ronda (2016) for a discussion on the relationship between maternal psychological distress and children's cognitive development.

This paper is closely related to Dahl and Lochner (2012) that exploits the exogenous variation of family income from the federal EITC expansions in the 1990s. They find that a $1000 increase in family income raises math and reading test scores by about 4% of a standard deviation.[10] The effects are larger for younger children (under age 12), for boys (doubled than for girls), and for black or Hispanic children. Dahl and Lochner (2012) also confirm that contemporaneous income has the largest effect on achievement. However, according to Nichols and Rothstein (2015), the interpretation of those estimates could be "hazardous" because the labor-supply effects may exact an independent role, in addition to the weak instrument issue.[11] Moreover, Dahl and Lochner (2012) do not address effects at early childhood stages, because most children in their sample (92%) are between ages 8 and 14, while children in PSID and its CDS in this paper can be as young as age 3 when developmental outcomes begin to be measured.

Along with other limited evidence linking EITC with children's educational outcomes, the empirical results are also conflicting. Most studies lack a theoretical model and their results vary by evaluation strategies based on different underlying identifying assumptions. One set of research has found positive effects concentrated among younger children (including Dahl and Lochner (2012), Michelmore (2013), Maxfield (2013)). The effects are then much smaller, or nonexistent, for older children, which

[10] In their original Dahl and Lochner (2012) paper, their result is around 6%. However, corrections have been made due to several coding error issues, and IV estimates falls by 32%, from 0.0610 to 0.0413, discussed in Lundstrom (2017) and Dahl and Lochner (2017)

[11] Weak instruments can lead to inconsistent and misleading results. See Bound et al. (1995); Stock et al. (2002a,b)

is consistent with the "window of opportunity" theory,. This theory emphasizes the importance of early childhood investment in shaping life-cycle skill formation. On the other hand, other studies (including Chetty et al. (2011), Manoli and Turner (2018), Bastian and Michelmore (2018)) have found only positive effects for older children in late adolescence, meaning around the time of high school graduation, which is consistent with the "cash on hand" hypothesis. Even certain *negative* effects have been found in Bastian and Michelmore (2018), where they find that family income from ages 6-12 has a negative impact (-0.0017) on a child's high school graduation, while family income from age 13-18 has a positive impact (0.0021). Both of these results are statistically significant at the same 10% level.[12] To mitigate such controversy, this paper recognizes the nature of children's human capital formation as a dynamic acquisition process, where at least three important channels are important (the goods input, the time input, and the non-parental child care). All of these inputs would be shifted differently by the EITC, so the overall effect not only depends on the productivity of each input but also the complementarity (or substitution) pattern across inputs. An ideal approach would be to estimate a production function of child skill formation in the spirit of Cunha and Heckman (2008) and Cunha et al. (2010), and to contain various inputs that are all affected by EITC policy variables. Unfortunately, this type of exercise is not the most feasible given the data structure as well as the sources of identification.

[12]See Table 6, Effect of EITC exposure on Education, Employment, and Earnings (IV Results) in Bastian and Michelmore (2018).

The rest of the paper proceeds as follows. Section 2.3 introduces the institutional background of the EITC in terms of eligibility and credit schedule. Section 2.4 contains a bare-bones model showing why the EITC may affect child development in which channels. Section 2.5 describes the data used in the empirical analysis and the construction of key variables. Section 2.6 presents the identification strategy based on EITC generosity. In Section 2.7 and Section 2.8, I provide the evidence of the EITC on different child developmental outcomes and potential mechanisms. Section 2.9 interprets the results and Section 2.10 concludes.

2.3 Institutional Background

Since introduced and enacted as part of the Tax Reduction Act of 1975, the EITC was made permanent in 1978 and experienced several major expansions with the Tax Reform Act (TRA) of 1986, the Omnibus Reconciliation Act (OBRA) of 1990, and the OBRA of 1993. Notably, OBRA 1993 was the largest among these expansions, doubling the payments to eligible families and sharply increasing the benefits to those with two or more children (**?**). In this section, I briefly introduce the eligibility and credit schedule of both federal and state EITC.

Eligibility. Eligibility for the federal EITC is based on two major criteria: (i) family structure, i.e., the number of "qualifying children" in the household, and (ii) family earnings and income. One of the main determinants for qualification is that a

child must be younger than 19 years old.[13] When the EITC was initially introduced, and until 1994, families without children could not benefit from it. Now, the benefit is much more generous for families with children and increases with the number of children. The second requirement is earned income, which must be positive and below a certain threshold. For families with positive tax liabilities from regular income tax or self-employment tax, the EITC is first used to offset these liabilities. Moreover, when the EITC exceeds other liabilities, it becomes refundable. In fact, over 85 percent of the EITC claimants benefit in terms of refunds.

Credit Schedule. The EITC schedule contains three different regions. The first region is called the "phase-in" or credit region, where the credit rate (*subsidy rate*) is positive (so the marginal tax rate, MTE, is negative). In this region, the higher a family's earning, the more tax credits can be claimed, until the earning reaches the threshold where the phase-in region ends. The second region is called the "flat" or "plateau" region, where a family earns the maximum amount of tax credit and the MTE becomes zero. This region ends as the family's income increases, finally reaching another threshold where the third region, "phase-out", starts. In the phase-out region, the EITC is reduced in proportion to the difference between earnings and the end of the plateau range, and the MTE is positive. The phase-out region ends at the next income threshold, when credit ultimately becomes zero and families are no longer eligible.

[13] 24 if a full-time student, or any age if totally disabled.

The expansion of the EITC is clearly illustrated in Figure 2.1, where the credit schedules for single parents with two children at 1979, 1993, 1996, and 2014 are depicted. As can be seen, there is moderate expansion from 1979 to 1993 in terms of both subsidy rate and the amount of maximum subsidy. Moreover, significantly increased generosity had taken place between 1993 to 1996, where the maximum subsidy nearly doubled and the subsidy rate increased by 100%. Though the phase-out rate had also been increasing over time, it was much flatter and grew more slowly than the subsidy rate.

State EITC. Since the late 1980s, some states have also begun to enact their own EITC schedules in addition to the federal EITC. The state-level EITC rule is much simpler, and is typically defined as a fixed percentage of the federal EITC. In addition to different percentages, another difference across states is whether the benefits are refundable. Detailed information about state-level EITC can be found at `http://users.nber.org/~taxsim/state-eitc.html`.

2.4 A Bare-Bones Model

To facilitate an understanding of the different mechanisms through which the EITC may affect child development, I propose a static model in this section to fix the

idea conceptually.

$$\begin{aligned} &\max_{C,H,MT,CC,M} U(C, L, K; \theta) \\ &K = f(\mu MT, M, CC) \\ &C + M + p^{CC} CC = WH + N + EITC \\ &L = TT - H - MT \end{aligned} \tag{2.1}$$

The utility function contains three components, where C is consumption, L is leisure, K is the child's human capital skill, θ is a vector of preference parameters varying in the population, and U is assumed to satisfy the usual concavity conditions. Leisure L is assumed to be a normal good.

The child's human capital skill is produced by different inputs, including maternal time (both quantity MT and quality μ), material goods M, and non-parental child care CC. Notably, child's human capital skill is not only affected by the amount of each output, but also by the underlying relationship across different inputs, i.e. whether inputs are *complementary* or *substitutable*.

The third line of Equation 2.1 constitutes the normal budget constraint, where the income comes from three components: (1) labor earnings WH as the product of wage rate W and hours of work H; (2) non-labor income N, including father's labor income for married parents[14]; and (3) the $EITC$.

The last line of Equation 2.1 captures the mother's time allocation, where her

[14]To save space, I choose not to incorporate father's labor income as a separate factor, that should then be augmented by an index function depending on whether the mother is married or single.

total time TT is allocated among three parts: (1) labor supply H, (2) child's time input MT , and (3) leisure L.

The $EITC$ can be further formulated as

$$EITC = \begin{cases} \psi^{in} \times WH & \text{if } WH \leq \Psi^{in} \\ \psi^{in} \times \Psi^{in} & \text{if } \Psi^{in} < WH \leq \Psi^{out} \\ \max\{0,\, \psi^{in} \times \Psi^{in} - \psi^{out} \times (WH - \Psi^{out})\} & \text{if } WH > \Psi^{out} \end{cases} \tag{2.2}$$

The EITC schedule produces a non-convex budget set that contains four regions and three kinks. Region I is the "phase-in" region, where the individual receives wage subsidy equaling ψ^{in} as long as the individual's income is less than the first income threshold Ψ^{in}. Region II is the "flat" region, where the individual receives the maximum credit $\psi^{in} \times \Psi^{in}$ as long as the individual's income is between the first and second income thresholds Ψ^{in} and Ψ^{out} respectively. Region II is the "phase-out" region, where the credit is reduced by rate ψ^{out} until the credit is fully phased out, when income is larger than the second threshold Ψ^{out}. Region IV is the original budget constraint without the EITC.

Based on this simple model, it is clear that the EITC can affect all three inputs of child's skill production function: goods, time, and child care.

First, the EITC leads to an eased budget constraint, which generates higher disposable income. This could potentially lead to higher material goods investment M, depending on the tradeoff between the marginal utility of the mother's own con-

sumption and the marginal utility of the chid's skill, which further depends on the productivity of goods investment compared to other inputs. If M is related to goods used to fulfill a child's daily nutritional and health needs or to books and studying resources, then this increase would be beneficial to the child. However, if the eased budget constraint does not directly result in increased M, this goods input channel may be more nuanced than it initially appears to be.

Second, the EITC will also change the mother's time allocation decision resulting from policy-induced labor supply change. On the one hand, theoretically speaking, the effect of the EITC on the time input MT is ambiguous because the effect on the mother's labor supply H is ambiguous, depending on which region she is in. On the other hand, the EITC has been mainly found to increase labor supply of single mothers at the extensive margin. The effect on married mothers has been found to be modestly negative at the intensive margin, while no effect has been observed at the extensive margin. Thus, it is important to discuss single mothers and married mothers separately. Moreover, it is unclear even for single mothers, whether an induced labor supply increase would lead to increased or decreased time investment in children, which is essentially a choice variable depending on both mother's observable and unobservable characteristics. A mother could choose to sacrifice her own leisure to both increase her labor supply and maintain the amount of time invested in her children. Moreover, MT only captures the change of quantity, while the quality of a mother's time μ can also be changed when the mother changes her labor force

participation decision. In the empirical analysis of this paper, maternal psychological distress is used as a proxy measure that could affect maternal time quality.[15]

Third, the maternal decision also depends on outside options, such as non-parental child care, the third input CC in child's skill production function. A working mother may need to choose alternative child care options, such as center-based child care or an informal arrangement with a neighbor or relatives, such as grandparents when she works full-time. In the case of single mother, an informal child-care arrangement might be a more likely option when she re-joins the labor force and starts to work. This could hinder child development if the quality of the informal child care is worse than the mother's own time input. Married mothers might be more likely to choose formal institutional child care programs, the quality of which is also heterogenous.

In sum, based on this simple model, I illustrate the various mechanisms through which the EITC may affect a child's human capital skill. The effect of the EITC on both time input and non-parental child care mechanisms is particularly ambiguous, not only due to the ambiguous effect of the EITC on maternal labor supply, but also because each behavioral response is a matter of her utility maximization.

[15]Conditional on the quantity of maternal time input, maternal psychological distress would affect the quality of the parental interaction with the child.

2.5 Data

2.5.1 The Panel Study of Income Dynamics

In this paper, I utilize the Panel Study of Income Dynamics (PSID) and its Child Development Supplement (CDS) and Time Diary (TD). In total, the PSID contains a representative sample of over 18,000 individuals living in 5,000 families in the United States.[16]

Beginning in 1997, the CDS collected information on 3,563 children aged 0 - 12 years old in 2,394 PSID families. A maximum of two children per family were randomly selected. Given that the children were between 0 and 12 years old in the first wave, their mothers were matched from PSID surveys in the years 1985 - 1997, as depicted in Table 2.17. The second and third waves of CDS were collected in 2002 and 2007, respectively. The CDS is particularly well-suited for this study because it provides detailed information about children's cognitive and non-cognitive outcomes, as well as the quantity of interaction between mother and child through its TD component [17].

Among the original 3,563 CDS children, there were 89 non-sample cases who

[16]The PSID is a nationally longitudinal study that began in 1968 with a core nationally representative sample called the Survey Research Center (SRC) sample, as well as an oversample that mainly represents low-income and black families, called the Survey of Economic Opportunity (SEO) sample.

[17]The unique feature of its Time Diary (TD) component makes PSID an ideal data set for studying the mechanisms through which the EITC affects child development. Other longitudinal micro-data, such as NLSY-79, has also been widely used in the context of child development, such as by Bernal (2008), Cunha and Heckman (2008), and Cunha et al. (2010). However, one of the major drawbacks is the lack of *quantity* of time investment in children.

were mistakenly selected for CDS, leading to an effective sample of 3,474 cases. The CDS children can be linked with the PSID core data, where maternal labor market outcomes, as well as other household characteristics are available (Table 2.13).

2.5.2 Variable Construction

In this section, I describe the measures of both child developmental outcomes and various mechanisms through which the EITC affects child development .

Children's Cognitive Skills. The CDS has used subtests from the Woodcock-Johnson Psych-Educational Battery-Revised (WJ-R) to objectively assess the children's academic skills in both reading and mathematics in all three waves of the survey.[18] Specifically, letter-word identification (LW) and passage comprehension (PC) tests are related to reading skills, and applied problems (AP) and calculations tests are related to mathematics skills. Descriptions of and the youngest age to participate in each test are summarized in Table 2.14.

Children's Non-cognitive Skills. The CDS contains two different types of measurements of children's non-cognitive skills. The first is the behavior problem index (BPI) which further includes two subscales: externalizing BPI and internalizing BPI. The former describes external or aggressive behaviors, while the latter describes internal or withdrawn behaviors. Both of these indexes have been commonly used in the literature, and the same sets of items have been used in other well-known longitudinal

[18] A detailed introduction can be found in Duffy and Sastry (2014)

studies as well, such as the National Longitudinal Survey of Youth (NLSY).[19] The second is the positive behaviors scale (PBS), which measures childhood emotional or social competence. Descriptions and the possible age range for participation are summarized in Table 2.15.

Maternal Time Investment. The availability of the time diary (TD), a unique feature of the CDS, provides detailed accounting of the type, number, duration, and location of children's activities during sampled 24-hour days, beginning at midnight and remaining as precise as possible on a minute-by-minute basis. One random weekday and one random weekend were chosen. Interview participants include 2,904 children, 2,569 children, and 1,442 children in the original CDS-I, CDS-II, and CDS-III, and the response rates were 82%, 88%, and 86%, respectively.[20]

In particular, the social context of each activity is specified, and a distinction is made between with whom the child was doing the activity and who else was present but not engaging. The former time a mother spent with her child was labeled as "engaged time" (also known as "active time" in the literature), while the latter was labeled as "accessible time" (also known as "passive time" in the literature). Both types of time investment could be productive as active engagement in shared activities has demonstrated to be important (Altintas (2016) and Kalil et al. (2012)), while passive accessible time input, or "being there", may also play a role in supervising,

[19] A detailed description can be found in The Panel Study of Income Dynamics Child Development Supplement User Guide Supplement: `https://psidonline.isr.umich.edu/cds/cdsi_usergd.pdf`

[20] The PCG completed the diary for the young children but older children and adolescents were expected to complete the time diaries themselves or with the assistance of the PCG, as needed.

monitoring, and facilitating children's activities. (Folbre et al. (2005), Sandberg and Hofferth (2001), and Sayer et al. (2004)).

Thus, three types of maternal time investment are constructed, "engaged", "accessible", and total (the sum of "engaged" and "accessible"), by taking a weighted sum $\frac{5}{7}$ for the weekday and $\frac{2}{7}$ for the weekend.

It is important to recognize that such a construction, though commonly used in the literature, does impose rigidity and may therefore amplify any atypicality.[21] This should be considered in later analysis. The TD also provides information about who filed the TD (mother alone, child alone, mother and child together, or others), whether it was completed on a weekday or weekend, and how typical the day was (from very typical to very atypical). These variables are therefore included in the later empirical analysis to alleviate the potential atypicality concern.

Maternal Child Care Arrangement. The CDS provides retrospective information about sample children's child care information in terms of type, frequency of use, and costs of arrangements for children through (and possibly after) kindergarten. When the CDS-I and CDS-II data are combined,[22] a near-complete picture of child care arrangement history from birth to the age of the child at the time of the CDS-II interview can be generated.

Thus, the retrospective information stored in CDS-I and CDS-II can be transformed into a longitudinal data structure, which is not only helpful to increase total

[21] CDS time diaries measure school year time.

[22] Children are aged 0 to 12 in CDS-I and 5 to 17 in CDS-II. In CDS-III, children will become at least age 10.

sample size, but significantly strengthens the identifying power, given that major federal EITC expansions took place in the early 1990s. Table 2.16 illustrates an example in which the specific calendar years have been listed between ages 0 and 5 for children in CDS-I. As discussed earlier in the institutional part of the EITC and as further discussed in greater detail in the next section, the three highlighted years (1986, 1990 and 1993) witnessed greatest federal EITC expansions, which have been used extensively in the difference-in-differences research design. Most children, meaning those older than age 3 in CDS-I, could experience at least one of such large federal expansion; for those who were aged between 7 and 12, they experienced two out of three expansions. By linking these child care arrangements to the maternal labor market outcomes in the PSID core data, we can at least partially utilize those federal expansions, to increase the power of identification.

I first construct a child age-specific binary dummy to indicate whether any non-parental child care had been used. I also explore different types of child care arrangements and generate two additional dummies to indicate whether the child care was done by any relative; or was from any formal program (such as Head Start, a pre-school program, or any other child care center). Though not perfect, these additional measures capture a certain degree of quality heterogeneity. Finally, I construct an age-specific weekly hours of child care arrangement, i.e., the intensive margin of non-parental child care usage.

Maternal Psychological Distress. Maternal psychological distress is another poten-

tial channel through which the EITC affects a child's skill formation. The PSID-CDS has provided the commonly used K-6 Non-Specific Psychological Distress Scale, developed by Ronald Kessler (Kessler et al. (2003)), to "discriminate cases of serious mental illness from non-cases in a general population survey". The K-6 includes six items answered by the primary caregivers about how they have been feeling during the prior four weeks. The original response items are based on a scale of 1-5, and the final K-6 scale has been recoded so that a lower score indicates a better condition. This measure is available in all three waves of PSID-CDS, as well as for the respondents in the PSID core data 2003.[23]

Others. Most variables associated with mothers' labor market outcomes as well as socioeconomic characteristics, come from the PSID core data, including the mother's age, race, years of education, state of residence, number of children, age of youngest child, annual earnings, and annual hours of work, among other data. Other child characteristics come from either the CDS or the PSID core data, including the child's gender, race, age, and other variables.

2.5.3 Summary Statistics

Henceforth, I make a distinction between two different analytical samples. The first one is called the *Original CDS*, which is based on three waves of PSID-CDS: 1997,

[23]The PSID has included the K-6 scale for all respondents from 2001 and onward, since the CDS-I, II, and III are from 1997, 2002, and 2007, respectively. The only additional information we can match from the PSID core data is thus 2003.

2002, and 2007. The second is called the *Augmented CDS* and contains the transformed longitudinal non-parental child care arrangements, as well as other household characteristics. It is worth noting that child development outcomes - both cognitive skills and non-cognitive skills - and maternal time input are only available in the Original CDS sample.

The summary statistics based on Original CDS data are displayed in Table 2.1. First, the EITC generosity, defined as the maximum federal and state EITC a family could possibly receive, indeed reveals large variation across the three years of CDS data. The smallest amount is about $3,000 while the largest amount exceeds $8,000. Approximately 20% of families in the CDS sample have ever received EITC. The percentage of African American mothers and children is about 40% due to the fact that the SEO component of the PSID has oversampled low-income families. In terms of the major input of the child skill production function, there is no significant difference between active time versus passive time, and approximately 40% of families have ever used child care arrangements, and the percentage of relative-based child care is almost double than that of the formal center-based child care.

To gain a understanding of the role of the EITC, I further split the sample, based on EITC eligibility as well as marital status, as documented in other columns of Table 2.1. On the one hand, I restrict the sample to EITC recipients and those without EITC, respectively. It seems that children in EITC families generally perform worse on both cognitive achievement tests and non-cognitive behavioral problems.

Mothers receiving EITC, on average, work less hours, earn less income, spend less time with their children, less frequently used formal child care (but use more relative-based child care), and have larger psychological distress issues. This initial glance is perhaps unsurprising given that the EITC is essentially part of the social safety net and is targeted toward low-income families. On the other hand, I also provide descriptive analysis by mother's marital status. Single mother families are much more likely to receive EITC and have worse socioeconomic outcomes and children's developmental outcomes when compared with married mother families.

2.6 Empirical Strategy

For the empirical analysis, I use *EITC generosity* (also known as EITC exposure) as the key variable of interest, which is defined as the maximum potential federal and state credit a child's family could receive given their state of residence, family size, and tax year, regardless of the family's income or parental marital status. Thus, this measure does not suffer from endogeneity, as would the actual amount of the EITC a family receives, which is determined by the choice of labor supply.

Theoretically, there are three main sources of exogenous variation: (1) the year the individual was born, which affects the generosity of the federal credit in a particular year; (2) the state of residence in each year, which leads to variations in state EITC generosity; and (3) the number of children in the household in each year, given the fact

that there are larger benefits for larger households. These three sources of variation lead to changes of the EITC generosity, which identifies the effect of EITC policy change on different variables of interest, including both child developmental outcomes and various mechanism variables. This approach can be interpreted as a continuous version of difference-in-differences method.

It is also worth noting that for results related to child developmental outcomes, given the timing of the CDS, i.e. 1997, 2002, and 2007, the identification cannot benefit much from the (1) variation[24] because the federal EITC schedule remained stable during that period, as illustrated in Figure 2.2. However, non-trivial (2) state EITC variation helps with identification, as depicted in Figure 2.3. The number of children could also ease the identification when families had more children from wave 1 to wave 2, e.g. having one child in CDS 1997 but two or more children in CDS 2002.[25]

When the Augmented CDS sample is used for the non-parental child-care arrangement, all (1), (2), and (3) sources of variation help with identification, particulary the (1), which is not possible with the Original CDS sample analysis. Major federal EITC expansions in late 1980s and early 1990s generate non-linear changes to EITC generosity, further strengthening the identification power.

[24](1) the year the individual was born, which affects the generosity of the federal credit in a particular year

[25]To better understand the different sources of variation, I conduct a simple analysis by regress the EITC generosity on year, number of children dummy, and child fixed effect. The results are displayed in Table 2.19 using the previously described PSID-CDS data. R^2 suggests how much proportion can be explained by these observable variations included in the regression, while $1 - R^2$ suggests how much can be explained by state heterogeneity.

Overall, I use three different empirical model specifications: (i) pooled OLS model with state fixed effect (Pooled OLS with State FE); (ii) child fixed-effect model (Child FE); and (iii) child value-added model (Child VA), each of which is discussed.

First, the pooled OLS model with state fixed effect is follows Equation 2.3:

$$Y_{ijst} = \alpha + \beta \text{EITC Generosity}_{jst} + \eta X_{ij} + \psi V_{st} + \lambda Z_s + \pi W_t + \epsilon_{ijst} \tag{2.3}$$

where i indexes individuals, j indexes families, s indexes states, and t indexes years. Y_{ijst} is the outcome variable of interest, which could either be child developmental outcomes or variables associated with different mechanisms. EITC Generosity$_{jst}$ represents the EITC generosity a family j could possibly receive given state s and year t. The coefficient of interest, β, represents the impact of an additional \$1,000 of EITC exposure on outcome variables of interest.

When the outcome variable of interest is child developmental outcomes, X_{ij} is a vector of personal characteristics that includes the child's gender, age, ethnicity,[26] and birth year; the mother's age, ethnicity, education, and marital status; and the number of children and age of the youngest child at the family level. When the outcome variable of interest is maternal time input, additional variables related to TD characteristics are also controlled, including the how typical the chosen weekday and weekend are, and who completed the TD, in addition to the variables previously mentioned in X_{ij}.

[26]Indicators for whether a child is African American, Hispanic, or other Non-White respectively.

V_{st} includes state-by-year policy and economic indicators, to alleviate the concern that state EITC may not be completely exogenous. Variables include GDP per-capita, personal income per-capita, unemployment rate, minimum wage, maximum welfare benefits for combined AFDC/TANF and SNAP, per-pupil K-12 public educational spending, and poverty rate. These variables have been commonly used in various contexts in the literature related to EITC. Z_s and W_t are state and year fixed effects, and ϵ_{ijst} is the idiosyncratic error term.

To add robustness, on the one hand, I conduct analysis by excluding high-income families and by focusing on low-educated versus high-educated families separately. On the other hand, in addition to the state fixed effect model, two additional empirical models are utilized: the child fixed effect model and the value-added model. The child fixed effect model is very similar to Equation 2.3, as illustrated in Equation 2.4:

$$Y_{ijst} = \alpha + \beta \text{EITC Generosity}_{jst} + \eta X_{ij} + \psi V_{st} + \lambda C_i + \pi W_t + \epsilon_{ijst} \tag{2.4}$$

where state fixed effect V_{st} is replaced with child fixed effect C_i. On the one hand, child fixed effect further eliminates child-specific time-invariant unobservables. On the other hand, the identification based on the child fixed effect model is slightly different from one based on the state fixed effect, when a family changes its state of residence. However, only a small portion of the sample changed their state of residence in the CDS years.

Finally, I also use a value-added model, augmented with a lagged outcome variable,

based on Todd and Wolpin (2003) and Todd and Wolpin (2007) and illustrated in Equation 2.5:

$$Y_{ijst} = \alpha + \beta \text{EITC Generosity}_{jst} + \rho Y_{ijs(t-5)} + \eta X_{ij} + \psi V_{st} + \lambda Z_s + \pi W_t + \epsilon_{ijst} \quad (2.5)$$

Since the lagged outcome variable is used as an additional control variable, the sample size of the analytic sample is reduced as expected. Thus, the interpretation associated with the value-added model is whether EITC generosity can explain the change of outcome variable values between two waves, or every 5 years in CDS.

2.7 EITC and Child Development

In this section, the effects of the EITC (as measured by EITC generosity) on various measures of a child's cognitive and non-cognitive skills is investigated separately.

2.7.1 Main Result

The baseline results based on pooled OLS with state fixed effect are documented in Table 2.2. Each column represents a different child development measure, where the first four columns are associated with child cognitive skills and the last three columns are associated with child non-cognitive skills. Table 2.2 also contains three panels. In Panel A, all CDS children are included in the analysis, while panels B and C distinguish by family structure: single-mother families only versus married-mother families only. For each child development measure, the estimated parameter of the

key variable of interest (EITC generosity) is presented, together with the mean of outcome variable ($\overline{Y}$), number of observations (N), and R-squared (R^2).

In Panel A of Table 2.2, with all CDS children included, different patterns can be observed between cognitive skills and non-cognitive skills. On the one hand, no effects of the EITC are specifically found on cognitive skills. On the other hand, the EITC has a significant effect on non-cognitive skills, with a particularly strong effect on reducing internalizing problems (0.079 S.D.). However, results differ by family structure.

As mentioned, Panel B of Table 2.2 only looks at children from single-mother families. A $1,000 increase in EITC generosity is found to be significantly related to a 0.08 to 0.11 S.D. decrease in reading skill across all three measures. A weaker negative effect is also found regarding mathematical skill. In contrast, the EITC leads to improved non-cognitive skills, such as reduced behavioral problems and increased positive behaviors. In particular, a $1,000 increase in EITC generosity significantly reduces internalizing behaviors by 0.13 S.D.. It is also important to note that children from single-mother families, on average, have worse cognitive and non-cognitive skills, as indicated by the below-zero and above-zero means of outcome variables, which is unsurprising given that single-mother families are often more socio-economically disadvantaged.

Panel C of Table 2.2 turns to married-mother families. The results are distinct from what are found among single-mother families. Namely, there is no such detri-

mental link between EITC and cognitive skills. The effects with respect to reading skills are neither economically meaningful (less than 0.01 S.D.) nor statistically significant. The EITC has a weak positive effect on mathematical skills. However, improved outcomes for children's non-cognitive skills are found. It is also worth noting that children from married-mother families, on average, have much better cognitive and non-cognitive skills, as indicated by the means of the outcome variables.

Overall, based on Table 2.2, notably distinct results have been found by family structure and by child skill type. On the one hand, the EITC is found to be detrimental to children's reading skills among those who come from single-mother families while no such effect are found for those who come from married-mother families. On the other hand, the EITC is found to improve children's non-cognitive skills, irrespective of family type.

Although the EITC generosity does not suffer from selection bias, one possible limitation underlies the pooled OLS with state fixed effect is it ignores the panel nature of the CDS data structure. To better account for the potential unobserved heterogeneity, and to test the robustness of our baseline results, I use child fixed effect model, the results of which are documented in Table 2.3. The overall structure of Table 2.3 is exactly the same as that of Table 2.2. Panel A again demonstrates that, overall, the EITC is not significantly linked with child cognitive skills but improves child non-cognitive skills. The latter positive effects are salient across all three non-cognitive measures: externalizing behavior problems, internalizing behavior problems,

and positive behavior. When the analysis is separated by family structure, the result again becomes distinct. As depicted in Panel B, children's reading test skills are found to be negatively linked with EITC generosity. However, the results are flipped when for children's non-cognitive skills, as the EITC reduces both externalizing and internalizing behavior problems. In Panel C, the EITC is not found to significantly affect children's cognitive skills among those from married-mother families. Meanwhile, the EITC also improves these children's non-cognitive skills, with a particular increase on the positive behavior scale.

It is also helpful to compare the results based on two specifications in Table 2.2 and Table 2.3. First, the pattern is highly consistent between the two tables across different panels. Second, the estimated effects tend to be larger under the fixed effect as compared to the pooled OLS with state fixed effect. For example, EITC generosity is found to reduce children's broad reading score by 0.11 S.D. under pooled OLS with state fixed effect, but the detrimental effect increases to 0.13 S.D. under child fixed effect. Such increase is also found generally regarding children's non-cognitive skills, irrespective of family structure.

Lastly, in addition to pooled OLS and child fixed effect models, a value-added model with lagged child outcome is also evaluated to further enhance the robustness, even though our empirical models do not suffer from endogeneity concerns associated with family inputs. Moreover, the value-added model is conceptually more appropriate because it is consistent with the argument of "self-productivity" of child skill

formation. The results are displayed in Table 2.4. First, there is a reduced number of observations, due to adding lagged outcome as an additional regressor. Thus, those children who only appear once in the CDS sample are left out of this analysis. Second, since CDS only has three waves, the first wave (CDS97) is essentially not used as an outcome variable in this specification. Thus, there lacks the opportunity to observe the effect of the EITC on younger children's developmental outcomes, a caveat worth considering when it comes to interpretation.

Overall, the estimated effects listed in Table 2.4 are smaller in magnitude, as expected, due to the emergence of lagged outcome as a regressor. However, the pattern remains exactly the same as in Table 2.2 and Table 2.3. Again, EITC has been found to have different impacts for different outcomes. Regarding cognitive skills, the EITC is found to be detrimental to child's reading scores, though the estimation becomes less precise due to a reduced sample size. For non-cognitive skills, the EITC is found to be beneficial, irrespective of family structure. Among various non-cognitive skills, internalizing behavior problems have the largest effect. Children from single-mother families benefit more than children from married-mother families.

2.7.2 Robustness

Several additional robustness checks are conducted. On the one hand, I exclude high-income households in the analytic sample. For example, in Table 2.20, 2.21,

and 2.22, I exclude families whose income are larger than $150,000, $100,000, and $85,000, respectively, since the EITC is only affecting low-income populations. The results turn out to be robust and the effects of the EITC remain significant and concentrated among single-mother families. It is worth noting that the estimated coefficients of interest are slightly larger. On the other hand, I also conduct the analysis separately by whether mothers have high school degree or not. Results for mothers who are high-school dropouts are depicted in Table 2.23 and for mothers who have at least high school degrees are depicted in Table 2.24. The results are also as expected, with more significant results and larger coefficients among mothers with lower education.

In addition, given the natural correlation between cognitive and non-cognitive skills, I also conduct the Seemingly Unrelated Regression (SUR) analysis where various child developmental outcomes are estimated jointly. In Table 2.30, each column represents the results among the full sample, single-mother families, and married-mother families respectively. The smaller standard errors come from the fact that SUR are more efficient due to the underlying correlation among outcome variables. The Breusch-Pagan test of independence is rejected at 1 percent level for each column. The correlation matrix of residuals also suggest that there are strong correlations between cognitive and non-cognitive skills.[27]

In sum, among different measures, the EITC is consistently found to have a nega-

[27]For example, the correlation of residuals between broad reading and internalizing behavioral problem is -0.14, while the correlation of residuals between internalizing and externalizing is 0.66 for the full sample. The full results of correlation matrix of residuals are available upon request.

tive impact on children's reading skills among those from single-mother families, while it benefits children's non-cognitive skills, especially reducing internalizing behavior problems, irrespective of family structure. This differs from Bastian and Michelmore (2018), who found a small positive effect of EITC generosity on children's composite test scores. However, their construction of composition have flaws, and their analysis did not distinguish between single-mother families and married-mother families.

2.8 Understanding Mechanisms

Given that the EITC has different impacts on children's cognitive and non-cognitive skills and such impacts differ by family structure, a natural question to ask, then, is why? If additional cash assistance coming from the credit is assumed to be helpful, or at least not harmful to child development, then there must be some additional maternal behavioral changes in response to the EITC that lead to the presented results. Understanding the underlying mechanisms through which the EITC generates heterogenous impacts is key to discussing meaningful policy implications.

In this section, I focus on the mechanisms suggested from economic theory: goods input, time input, and non-parental child care input. Contributions are made in all three mechanisms, with a particular emphasis on the latter two. Moreover, I attempt to go beyond the quantity to investigate the potential quality heterogeneity. For the former, I explore the impact of the EITC on maternal psychological distress,

which may affect the quality of maternal time with the child. For the latter, I explore different types of child care and distinguish between care given by relatives and institutional child care programs.

2.8.1 Goods Input Channel

In Table 2.5, it is first verified whether EITC generosity is associated with an increase in the imputed amount of EITC eligibility and family resources. As expected, EITC generosity is found to be predictive of imputed amount of EITC generosity (column 1 of Table 2.5). In Panel A, an $1,000 of EITC generosity is linked with $180 of EITC eligibility, which is fairly consistent with what has been found in the literature. For example, Bastian and Michelmore (2018) obtained an estimate of $160 also using PSID. Overall, EITC generosity also predicts an increased total family income, as $1,000 of EITC generosity is linked with $1,270 total family income. In Panel B of Table 2.5, single-mother families exhibit a much larger linkage between EITC generosity and EITC eligibility, while the link is almost negligible for married-mother families (Panel C). Moreover, the linkage between EITC generosity and total family income is weaker for married-mother families.

Moreover, there is both qualitative (Edin and Lein (1997)) and quantitative (Meyer and Sullivan (2008)) evidence suggesting that income differs from consumption among single-mother families, both promoting the use of household consumption expenditure instead of household income or wealth to better measure children's ma-

terial well-being. As Hao and Yeung (2015) argue, distinguishing between household income and household consumption expenditure helps conceptually clarify that household income measures resources, while household consumption expenditure measures the results of resource allocation.[28]

In the second and third waves of the CDS, several questions associated with child-specific expenditures are asked. Though these questions cannot provide a complete measure of parental goods input in children. They can at least provide some clues. I construct an aggregate measure of child-specific expenditure that covers four categories commonly existing in two waves: clothes, toys, school supplies, and travel.[29] The results are displayed in column 3 of Table 2.5. No significant link is found between EITC generosity and child expenditure, irrespective of family structure. However, the interpretation warrants caution, since there are only two waves of expenditure data, which may also suffer from measurement error and lack sufficient identifying power resulting from EITC generosity variation. Nevertheless, the exercise here at least illustrates that the transmission from family-level total income to child-level expenditure might be less clear and more complex than a linear mapping.

In Table 2.6, child fixed effect confirms what has previously been found, though

[28]There is also potentially a race difference, i.e., the black-white gap. Charles et al. (2009) found that blacks spent significantly more on "visible goods" (such as clothes, cars, and jewelry) than whites with comparable income, in an effort to seek social status in communities. Meanwhile, blacks tend to spend a lower share of their expenditures on education than did their white counterparts (Fan and Lewis (1999)). However, Hao and Yeung (2015) found the opposite, i.e., black parents spend less on status-signaling items than did comparable white parents, particularly in the lower-expenditure groups.

[29]Child-specific food expenditure is also asked but only in the second wave.

it likely suffers more from attenuation bias because both income and expenditure are self-reported.

2.8.2 Time Input Channel

Time input could be particularly salient in the context of the EITC, which has been found to shift maternal labor supply decisions. The unique feature of the TD module of CDS allows one to precisely examine if maternal time input might be affected by EITC. Moreover, I am able to not only look at the total quantity of maternal time input, but also differentiate between two types: the engaged maternal time input and the accessible maternal time input.

In Table 2.7, I first provide the overall effect without splitting the sample (Panel A). On the one hand, a negative effect of EITC is revealed in the first column on maternal engaged time. A $1,000 increase in EITC generosity is associated with a 0.55 weekly hour reduction in maternal engaged time. On the other hand, a positive effect of EITC has been found on maternal accessible time - a $1,000 increase in EITC generosity is associated with a 1.28 weekly hour increase in maternal accessible time. This shift between the two different types helps explain the insignificant result found in the third column, with respect to total weekly hours spent with child. Given that, on average, mothers spend 17 hours per week in engaged time and accessible time, these effects represent about 3 percent and 7 percent reduction and increase, respectively.

When dividing the sample by family structure, distinct patterns emerge again. In Panel B, single-mothers are significantly reducing their engaged time (1.69 hours per week) at 1 percent level in response to $1,000 EITC generosity increase. This maps to an 11 percent reduction, given that, on average, single-mothers spend about 15 hours per week in engaged time. The number is large in magnitude. In contrast, single-mothers also tend to increase their accessible time: 1.76 hours per week maps to an 11 percent increase. Overall, there is a slight reduction in total time single-mother spend with their children, though it is not precisely estimated. In Panel C, EITC is not found to be negatively related to any time input category, which is distinct from what has been found associated with single-mother families.

Such finding is qualitatively similar when compared to results in child fixed effect model (Table 2.8), even though the coefficients are less precisely estimated.

2.8.3 Non-Parental Child Care Channel

Third, I look at the non-parental child-care channel. It is worth mentioning that the data used in this section is no longer the three waves of CDS data. Instead, the augmented CDS sample, a constructed panel data based on retrospective child-care history in CDS 1997 and CDS 2002 is used, in addition to the original CDS. The advantage is clear: to take advantage of major federal EITC expansions since the fiscal year can be as early as 1983.

The patterns between Table 2.9 and Table 2.10 are similar and I focus the dis-

cussion on Table 2.10, since it is more precisely estimated. Overall (Panel A), it suggests that the EITC leads to increased child care usage at both extensive and intensive margins and across two different types, i.e. by relative and formal child care programs.

As depicted in Panel B, single-mother families are found to increase any child care usage by 5.8 percentage point, which maps to about a 15 percent increase. However, when considering different types, it is striking to observe that single mothers are mainly choosing informal child care by relatives, which is commonly considered to be of low quality. The magnitude is large, since 3.7 percentage point increase in child care by relatives maps to a 20 percent increase. In contrast, there is no significant impact of EITC on formal child care programs. The effect on intensive margin is also large: 4.10 weekly hours of increase in any child care program maps to over 20 percent increase.

In Panel C, married mother families share similar results in terms of increased child care usage at both the intensive and extensive margin. However, one major distinction lies in the child care type. Unlike single-mother families who increase child care by relatives, married-mother families increase formal child care usage instead.

Though it may be difficult to compare the quality of formal child care centers versus maternal time for married mothers, it is arguably fair to infer that child care by relatives is detrimental to children of single-mother families. This may further help explain the fact that the EITC leads to worse reading skills for children from

single-mother families, together with the evidence that single-mothers tend to shift from engaged time to accessible time.

2.8.4 Psychological Distress

As demonstrated in the previous section, single mothers are found to respond actively to the EITC, i.e. reduce their engaged time input while increasing their accessible time input with their children. This is more about the quantity perspective of time. Another perspective is related to quality of time. If the quality of maternal time were increased, even with reduced quantity of time, a child would not necessarily be worse off. In contrast, if both quality and quantity of time were decreased, then a child could become worse off.

The maternal psychological distress is used as one proxy measure for the quality of maternal time. Tables 2.11 and 2.12 suggest that single-mothers would be more likely to become distressed, while married mothers would possibly become less distressed. The former effect is significantly estimated at the 1 percent level using the child fixed effect model in Panel B of Table 2.12, which represents a nearly 20 percent increase. The latter effect is negative, meaning a reduction in maternal psychological distress, though it is not precisely estimated.

The evidence suggests that not only the quantity, but also the quality, of maternal time input, might be responsible for explaining the negative impact on children's cognitive skills among single-mother families.

2.9 Interpreting Empirical Results

Being equipped with an understanding of the various mechanisms discussed, an ideal next step is to estimate a production function of child skill formation in the spirit of Cunha and Heckman (2008) and Cunha et al. (2010), which contains various inputs, all inputs instrumented with EITC policy variables. Assuming this can be done, then one would be able to interpret the results in terms of how the EITC leads to specific changes of inputs in the child skill production function, which ultimately results in corresponding changes in child skill formation. Unfortunately, this type of exercise is hardly feasible given the data structure I have, as well as the sources of identification, which requires EITC policy-related exclusion restrictions or instrumental variables for each of the inputs, especially for different types of maternal time and for non-parental child care.

Although not directly possible, I try to indirectly interpret the results based on results from other literature related to these mechanisms and how they shape child skill formation.

First, the literature has demonstrated that engaged and enriching maternal time is particularly significant to child development in the U.S. (Del Boca et al. (2014)), Australia (Fiorini and Keane (2014)), and the UK (Del Bono et al. (2016)). On the one hand, a reduction in engaged time could be detrimental to a child's cognitive development. On the other hand, a shift from engaged time to accessible time and keeping the total time roughly unchanged could be less harmful to a child's non-

cognitive skills.

Second, there is strong evidence that center-based former child care can generate better child outcomes than at-home informal child care (Heckman et al. (2013), Kline and Walters (2016), and Feller et al. (2016)). Bernal (2008); Bernal and Keane (2010, 2011) have also found that non-parental child care is detrimental to children's cognitive development for both single mothers and married mothers, even though they have not differentiated much at the quality dimension. A detrimental effect of non-parental child care on preschool-aged children's cognitive development is also found by Herbst (2013). Therefore, the fact that both single mothers and married mothers increase non-parental child care usage but single mothers are more likely to choose informal child care by relatives while married mothers are more likely to choose formal, center-based child care programs, provides a rationale for the detrimental effect of EITC generosity on the cognitive skills of children who come from single-mother families.

Third, Ronda (2016) illustrates the different mechanisms through which maternal psychological distress may impair a child's cognitive development. EITC generosity leads to increased psychological distress for single mothers, which may consequently result in negative effects on children's cognitive development. Herbst (2017) also finds that working mothers experience an increase in depressive symptoms as a mechanism through which the welfare reform's work requirements negatively affect child welfare, which shares a similar perspective to what has been found in the present research.

This could be possible, given the pro-working nature of the EITC that generates strong labor supply incentive for mothers which is similar to the welfare reform's work requirements.

Last, much less is known about determinants of children's non-cognitive skills, aside from the recent work of Akee et al. (2018), which highlights the importance of family income in affecting personality traits. If the positive effect of family income is stronger for non-cognitive skills than for cognitive skills, or if the total maternal time with a child matters more than the specific-type of time in affecting non-cognitive skills as compared with cognitive skills, then the varying results between cognitive and non-cognitive skills could be partially resolved.

2.10 Discussions and Conclusions

This paper investigates the mechanisms through which the EITC affects maternal decision making and child skill formation using the PSID and its CDS.

On the one hand, among different measures, the EITC has been found to have a negative impact on children's cognitive skills among those come from single-mother families, as compared to their counterparts from married-mother families. On the other hand, the EITC is associated with improvements in children's non-cognitive skills. The positive effect is salient for children from both single-mother families and married-mother families. Moreover, the effect is much larger for the more disadvan-

taged children, which differs from the case of cognitive skills. The results are robust based on different empirical model specifications.

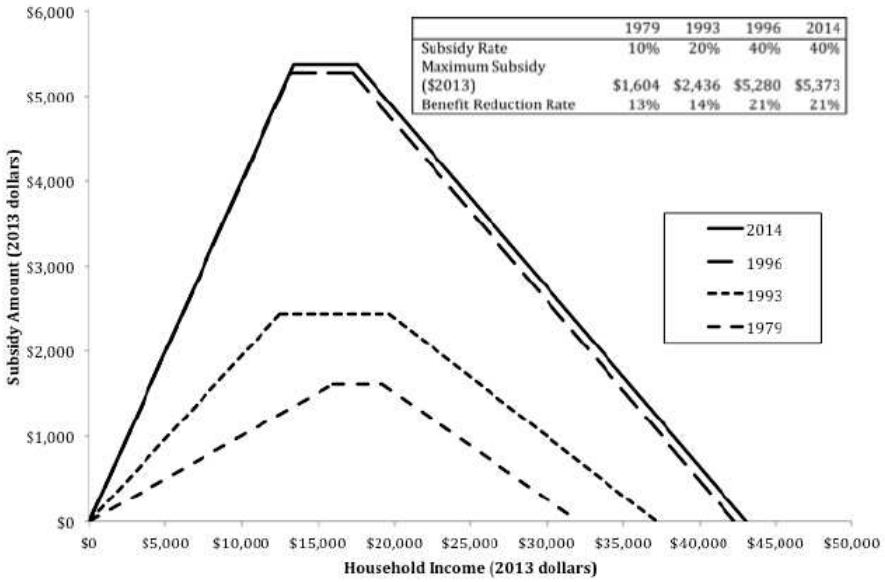
At least three channels may help explain the effect of the EITC on child outcomes. The most striking channel comes from non-parental child care. Both single mothers and married mothers tend to their increase child care usage at both extensive and intensive margins. However, single mothers are more likely to choose informal child care arrangements by relatives, which is usually thought to be of low quality and may have negative consequences for child development. In contrast, married mothers are more likely to use formal child care. Second, single mothers are likely to respond by changing their time inputs by switching between time types, i.e. reducing their actively engaged time but increasing accessible time. Such a shift is not observed among married mothers. Reduced engaged time may also be detrimental to children's cognitive development. Third, single mothers are worse off and more distressed when faced with increased EITC generosity. However, the EITC has a modest positive effect on married mother's psychological distress. The combined effects based on these three channels may help explain why children from single-mother families exhibit worse cognitive skills. Finally, whether increased family income directly leads to increased child-specific expenditures remains unclear and requires better data; this should therefore be explored in future research.

These findings also shed light on potentially important policy implications. One of the most direct policy suggestions is to promote the accessibility of high-quality child

care for disadvantaged families. As summarized by Hotz and Wiswall (2019), the current policy landscape on child care policies also has its limitations. For example, child care subsidies have consistently shown that subsidy receipt reduces cognitive development and increases behavioral problems (Herbst and Tekin (2010, 2016)). Head Start may be different, the heterogeneous effect of which, as well as the importance of care substitution, has been emphasized recently. (Feller et al. (2016); Kline and Walters (2016); Morris et al. (2018)). One of the most relevant findings is that the impacts of Head Start are substantially larger for children who would have otherwise received parental or informal non-parental care and are small or non-existent for children who would have otherwise been cared for via center-based formal centers. Thus, promoting Head Start for children from single-mothers families in the context of the EITC could potentially be beneficial, as these mothers have been found to increase their usage of informal child-care arrangements by relatives. A specific comparison of how different child care policies may differently alleviate the detrimental effects revealed in this paper in the context of the EITC is beyond the scope of this paper and remains to be explored in future research.

2.11 Figures

Figure 2.1: EITC Schedule for Single Parents with Two Qualifying Children



Note: 1979,1993, 1996 and 2014; obtained from Nichols and Rothstein (2015); calculations assume no unearned income.

Figure 2.2: Federal EITC Evolution

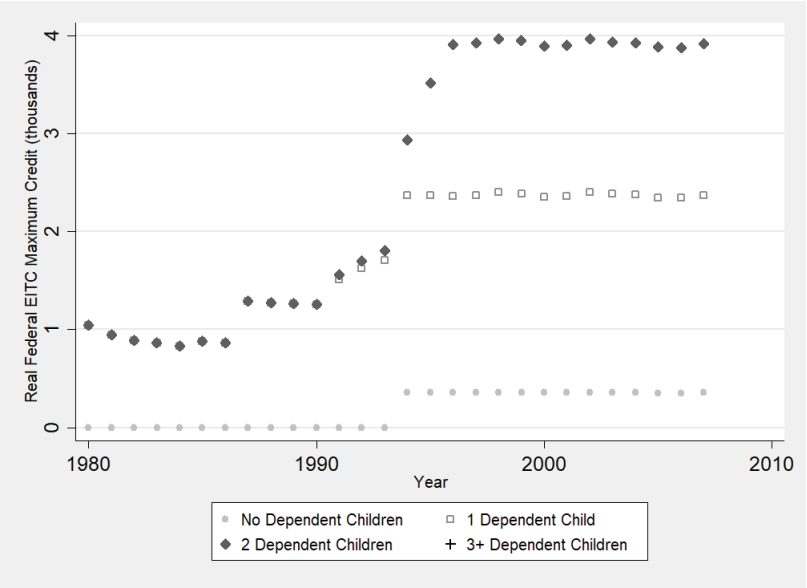
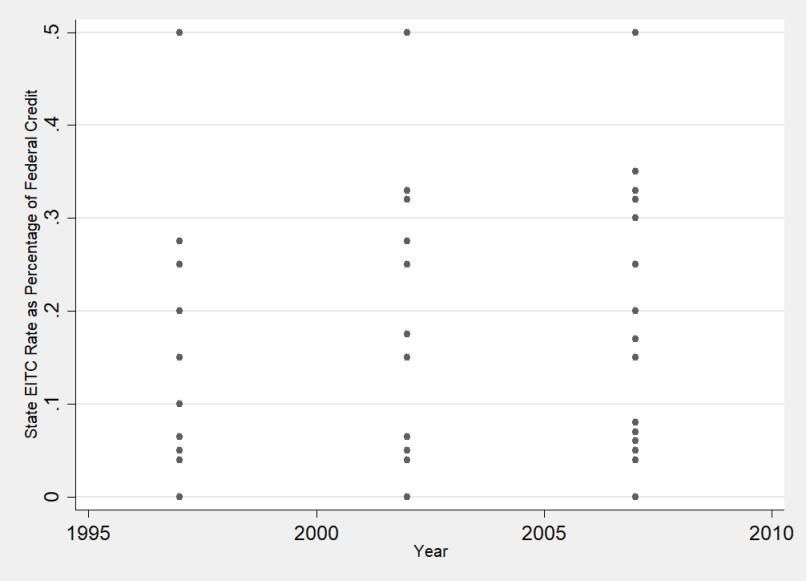


Figure 2.3: State EITC Evolution



2.12 Tables

Table 2.1: Summary Statistics Based on CDS Data

	Overall		with EITC		w/o EITC		Single		Married	
	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd
EITC Generosity	3.73	0.76	3.77	0.71	3.71	0.78	3.71	0.82	3.73	0.73
Household: Whether Has EITC or Not	0.31	0.46	1.00	0.00	0.00	0.00	0.60	0.49	0.18	0.38
Household: $ of Federal EITC	0.56	1.06	1.82	1.18	0.00	0.00	1.16	1.32	0.29	0.79
Household: $ of State EITC	0.02	0.13	0.06	0.24	0.00	0.00	0.04	0.21	0.01	0.07
Household: $ of Total EITC	0.58	1.11	1.88	1.24	0.00	0.00	1.21	1.39	0.30	0.81
Letter Word Z Score	0.03	1.00	-0.27	0.93	0.15	1.00	-0.28	0.97	0.17	0.97
Passage Comprehension Z Score	0.03	1.00	-0.27	0.93	0.16	1.00	-0.31	0.98	0.19	0.97
Broad Reading Z Score	0.03	1.00	-0.30	0.92	0.17	1.00	-0.32	0.99	0.19	0.96
Applied Problems Z Score	0.07	0.98	-0.28	0.89	0.22	0.98	-0.26	0.94	0.22	0.97
BPI: Externalizing Z Score	-0.04	0.98	0.09	1.01	-0.09	0.97	0.13	1.03	-0.12	0.95
BPI: Internalizing Z Score	-0.03	0.97	0.04	1.00	-0.06	0.96	0.05	1.04	-0.07	0.94
PBS: Z Score	0.03	0.98	0.04	1.03	0.02	0.96	-0.01	1.02	0.05	0.96
Child: Girl	0.50	0.50	0.50	0.50	0.49	0.50	0.51	0.50	0.49	0.50
Child: Black	0.39	0.49	0.55	0.50	0.32	0.47	0.72	0.45	0.24	0.43
Child: Hispanic	0.08	0.26	0.13	0.34	0.05	0.22	0.04	0.18	0.09	0.29
Child: Age	10.35	4.69	10.07	4.64	10.48	4.70	10.85	4.48	10.13	4.76
Mother: Black	0.38	0.49	0.55	0.50	0.31	0.46	0.71	0.46	0.24	0.43
Mother: Unmarried at Child Birth	0.33	0.47	0.54	0.50	0.23	0.42	0.68	0.47	0.17	0.38
Mother's Age at Child Birth	26.45	5.84	24.39	5.81	27.37	5.62	24.73	6.12	27.22	5.54
Mother: Age	36.31	7.21	33.97	7.19	37.35	6.97	35.09	7.59	36.85	6.96
Mother: Married	0.69	0.46	0.40	0.49	0.82	0.38	0.00	0.00	1.00	0.00
Mother: Years Completed Education	12.95	2.38	11.85	2.26	13.44	2.27	12.35	1.97	13.22	2.50
Mother: Cognitive Skill	30.87	5.39	28.51	5.16	31.92	5.15	28.48	5.41	31.90	5.04
Household: Number of Children	2.30	1.06	2.43	1.08	2.24	1.05	2.32	1.14	2.29	1.03
Household: Age of Youngest Child	6.84	4.47	6.10	4.24	7.17	4.53	6.97	4.45	6.78	4.48
Mother: No Work	0.21	0.41	0.18	0.38	0.23	0.42	0.19	0.39	0.22	0.42
Mother: Annual Hours of Work	1269.30	940.33	1212.01	885.30	1294.64	962.67	1382.81	942.66	1220.39	935.19
Household: Total Income with EITC	50.08	31.69	25.24	12.97	61.94	31.12	27.60	19.53	60.21	30.90
Child Expenditure	1.06	0.72	0.80	0.61	1.18	0.74	0.88	0.67	1.15	0.73
Mother: Weekly Hours of Active Time	16.94	11.28	16.25	11.26	17.23	11.28	14.91	11.24	17.78	11.20
Mother: Weekly Hours of Passive Time	16.73	11.15	16.28	11.18	16.91	11.13	15.42	11.27	17.27	11.06
Mother: Weekly Hours of Total Time	35.90	15.92	35.18	15.99	36.20	15.88	33.18	15.85	37.01	15.81
Mother: Whether Uses Any CC	0.31	0.46	0.34	0.47	0.30	0.46	0.38	0.48	0.29	0.45
Mother: Whether Uses CC by Relatives	0.16	0.37	0.21	0.41	0.14	0.35	0.23	0.42	0.13	0.34
Mother: Whether Uses Formal CC	0.12	0.32	0.11	0.31	0.12	0.33	0.13	0.33	0.11	0.32
Mother: Weekly Hours of CC	5.75	12.33	6.61	13.09	5.37	11.97	7.13	13.49	5.16	11.75
Mother: Psychological Distress Index	3.93	3.62	4.66	4.00	3.62	3.39	4.74	4.19	3.57	3.27

Note: (1) Number of observations for the overall sample is 6915. It is 2126, 4789, 2118, and 4797 for households with EITC, households without EITC, households with single-mothers, and households with married-mothers respectively. (2) Household EITC-related variables (whether has EITC or not, actual amount of federal EITC, actual amount of state EITC, and actual amount of total EITC) are simulated using relevant information in PSID and NBER's TAXSIM. (3) Actual amount of federal/state/total EITC, total income with EITC are scaled in $1,000. All monetary variables are in 2014 real values.

Table 2.2: Impact of EITC on Child Developmental Outcomes

(Original CDS - Full Sample, Pooled OLS with State FE)

	Cognitive Skill				**Non-cognitive Skill**		
	Letter Word	Passage Comprehension	Broad Reading	Applied Problem	Externalizing Problem	Internalizing Problem	Positive Behavior
				Panel A: Full CDS Families			
EITC Generosity	-0.022	-0.031	-0.031	0.012	-0.057*	-0.079***	0.025
	(0.038)	(0.025)	(0.029)	(0.030)	(0.032)	(0.028)	(0.028)
$\overline{Y}$	0.035	0.050	0.044	0.087	-0.044	-0.025	0.012
N	4296	3762	3760	3815	4465	4474	4513
R^2	0.226	0.263	0.280	0.289	0.068	0.061	0.050
				Panel B: Single - Mother Families Only			
EITC Generosity	-0.085*	-0.090**	-0.11**	-0.054	-0.057	-0.13**	0.0083
	(0.050)	(0.044)	(0.050)	(0.053)	(0.052)	(0.050)	(0.049)
$\overline{Y}$	-0.315	-0.330	-0.350	-0.270	0.141	0.043	-0.020
N	1311	1159	1158	1168	1365	1367	1382
R^2	0.176	0.222	0.228	0.222	0.077	0.117	0.085
				Panel C: Married - Mother Families Only			
EITC Generosity	0.0018	-0.018	-0.0057	0.040	-0.073**	-0.056	0.038
	(0.040)	(0.032)	(0.032)	(0.036)	(0.036)	(0.034)	(0.036)
$\overline{Y}$	0.189	0.220	0.220	0.244	-0.126	-0.055	0.026
N	2985	2603	2602	2647	3100	3107	3131
R^2	0.198	0.227	0.244	0.270	0.059	0.057	0.051

Note: Standard errors in parentheses and are clustered at the state-level; * p<0.1, ** p<0.05, *** p<0.01. Each cell represents a separate regression with full set of control variables, including child's gender, race, age, age square; mother's race, age, age square, marital status, educational attainment, whether mother is unmarried at child birth; number of children and age of the youngest child in the household; region, state, and year dummy; state policy and characteristics including per-capital real GDP, unemployment rate, real state minimum wage, combined state and federal current per pupil real spending on K-12 public education, combined monthly maximum AFDC/TANF and Food Stamps benefits for 3-person family. The full results including coefficients for all variables are available upon request.

Table 2.3: Impact of EITC on Child Developmental Outcome

(Original CDS - Full Sample, Child FE)

	Cognitive Skill				**Non-cognitive Skill**		
	Letter Word	Passage Comprehension	Broad Reading	Applied Problem	Externalizing Problem	Internalizing Problem	Positive Behavior
				Panel A: Full CDS Families			
Full Sample	-0.046	-0.071*	-0.042	-0.0076	-0.074**	-0.10***	0.082**
	(0.034)	(0.040)	(0.033)	(0.033)	(0.034)	(0.035)	(0.039)
$\overline{Y}$	0.035	0.050	0.044	0.087	-0.044	-0.025	0.012
N	4296	3762	3760	3815	4465	4474	4513
R^2	0.050	0.057	0.052	0.113	0.022	0.042	0.012
				Panel B: Single-Mother Families Only			
Single-Mother Sample	-0.11*	-0.13	-0.13**	-0.0024	-0.22***	-0.20**	0.090
	(0.064)	(0.083)	(0.064)	(0.086)	(0.084)	(0.086)	(0.11)
$\overline{Y}$	-0.315	-0.330	-0.350	-0.270	0.141	0.043	-0.020
N	1311	1159	1158	1168	1365	1367	1382
R^2	0.070	0.118	0.114	0.167	0.038	0.082	0.041
				Panel C: Married-Mother Families Only			
Married-Mother Sample	-0.021	-0.050	-0.013	0.011	-0.037	-0.067	0.11***
	(0.045)	(0.055)	(0.045)	(0.044)	(0.041)	(0.042)	(0.043)
$\overline{Y}$	0.189	0.220	0.220	0.244	-0.126	-0.055	0.026
N	2985	2603	2602	2647	3100	3107	3131
R^2	0.068	0.059	0.069	0.113	0.028	0.042	0.017

Note: Standard errors in parentheses and are clustered at the child-level; * p<0.1, ** p<0.05, *** p<0.01. Each cell represents a separate regression with full set of control variables, including child's gender, race, age, age square; mother's race, age, age square, marital status, educational attainment, whether mother is unmarried at child birth; number of children and age of the youngest child in the household; year dummy; state policy and characteristics including per-capital real GDP, unemployment rate, real state minimum wage, combined state and federal current per pupil real spending on K-12 public education, combined monthly maximum AFDC/TANF and Food Stamps benefits for 3-person family. The full results including coefficients for all variables are available upon request.

Table 2.4: Impact of EITC on Child Developmental Outcome

(Original CDS - Full Sample, Child VA)

	Cognitive Skill				**Non-cognitive Skill**		
	Letter Word	Passage Comprehension	Broad Reading	Applied Problem	Externalizing Problem	Internalizing Problem	Positive Behavior
				Panel A: Full CDS Families			
Full Sample	-0.034	-0.067*	-0.065**	-0.013	-0.077**	-0.12***	0.064*
	(0.031)	(0.036)	(0.033)	(0.030)	(0.032)	(0.033)	(0.033)
$\overline{Y}$	0.022	-0.031	0.001	0.022	-0.076	-0.028	0.006
N	2085	1652	1650	1697	2229	2229	2266
				Panel B: Single-Mother Families Only			
Single-Mother Sample	-0.063	-0.100*	-0.097*	-0.068	-0.11*	-0.16***	0.0099
	(0.052)	(0.056)	(0.050)	(0.050)	(0.062)	(0.059)	(0.059)
$\overline{Y}$	-0.325	-0.410	-0.393	-0.362	0.120	-0.020	-0.011
N	652	513	512	531	696	693	710
				Panel C: Married-Mother Families Only			
Married-Mother Sample	-0.028	-0.048	-0.055	0.0036	-0.055	-0.10**	0.089**
	(0.041)	(0.047)	(0.043)	(0.039)	(0.037)	(0.040)	(0.038)
$\overline{Y}$	0.180	0.140	0.179	0.197	-0.165	-0.031	0.014
N	1433	1139	1138	1166	1533	1536	1556

Note: Standard errors in parentheses and are clustered at the child-level; * p<0.1, ** p<0.05, *** p<0.01. Each cell represents a separate regression with full set of control variables, including child's gender, race, age, age square; mother's race, age, age square, marital status, educational attainment, whether mother is unmarried at child birth; number of children and age of the youngest child in the household; region, state, and year dummy; state policy and characteristics including per-capital real GDP, unemployment rate, real state minimum wage, combined state and federal current per pupil real spending on K-12 public education, combined monthly maximum AFDC/TANF and Food Stamps benefits for 3-person family. The full results including coefficients for all variables are available upon request.

Table 2.5: Impact of EITC on Goods Input

(Original CDS - Full Sample, Pooled OLS with State FE)

	EITC Eligibility	Total Income with EITC	Child-Expenditure
	Panel A: Full CDS Families		
Full Sample	0.18***	1.27**	-0.040
	(0.028)	(0.54)	(0.027)
$\overline{Y}$	0.578	50.159	1.067
N	6262	5797	2853
R^2	0.237	0.488	0.171
	Panel B: Single-Mother Families Only		
Single-Mother Sample	0.39***	0.52	-0.048
	(0.046)	(0.71)	(0.052)
$\overline{Y}$	1.246	26.526	0.876
N	1896	1778	943
R^2	0.121	0.274	0.143
	Panel C: Married-Mother Families Only		
Married-Mother Sample	0.061*	1.66*	-0.026
	(0.035)	(0.86)	(0.033)
$\overline{Y}$	0.288	60.614	1.162
N	4366	4019	1910
R^2	0.153	0.350	0.157

Note: Standard errors in parentheses and are clustered at the child-level; $*p < 0.10, **p < 0.05, ***p < 0.010$. Each cell represents a separate regression with full set of control variables, including child's gender, race, age, age square; mother's race, age, age square, marital status, educational attainment, whether mother is unmarried at child birth; number of children and age of the youngest child in the household; region, state, and year dummy; state policy and characteristics including per-capital real GDP, unemployment rate, real state minimum wage, combined state and federal current per pupil real spending on K-12 public education, combined monthly maximum AFDC/TANF and Food Stamps benefits for 3-person family. The full results including coefficients for all variables are in the online appendix or available upon request.

Table 2.6: Impact of EITC on Goods Input

(Original CDS - Full Sample, Child FE)

	EITC Eligibility	Total Income with EITC	Child-Expenditure
	Panel A: Full CDS Families		
Full Sample	0.084**	0.94	-0.099**
	(0.033)	(0.78)	(0.042)
$\overline{Y}$	0.578	50.159	1.067
N	6262	5797	2853
R^2	0.086	0.234	0.080
	Panel B: Single-Mother Families Only		
Single-Mother Sample	0.20*	1.98	-0.036
	(0.10)	(1.24)	(0.075)
$\overline{Y}$	1.246	26.526	0.876
N	1896	1778	943
R^2	0.104	0.170	0.154
	Panel C: Married-Mother Families Only		
Married-Mother Sample	0.045	1.04	-0.16***
	(0.029)	(1.02)	(0.054)
$\overline{Y}$	0.288	60.614	1.162
N	4366	4019	1910
R^2	0.041	0.226	0.110

Note: Standard errors in parentheses and are clustered at the child-level; $*p < 0.10, **p < 0.05, ***p < 0.010$. Each cell represents a separate regression with full set of control variables, including child's gender, race, age, age square; mother's race, age, age square, marital status, educational attainment, whether mother is unmarried at child birth; number of children and age of the youngest child in the household; region, state, and year dummy; state policy and characteristics including per-capital real GDP, unemployment rate, real state minimum wage, combined state and federal current per pupil real spending on K-12 public education, combined monthly maximum AFDC/TANF and Food Stamps benefits for 3-person family. The full results including coefficients for all variables are in the online appendix or available upon request.

Table 2.7: Impact of EITC on Maternal Time Input

(Original CDS - Full Sample, Pooled OLS with State FE)

	Engaged Time	Accessible Time	Total Time
	Panel A: Full CDS Families		
Full Sample	-0.55	1.27***	0.40
	(0.38)	(0.37)	(0.42)
$\overline{Y}$	17.316	17.133	36.416
N	4223	4205	4135
R^2	0.241	0.067	0.183
	Panel B: Single-Mother Families Only		
Single-Mother Sample	-1.69***	1.76**	-0.20
	(0.55)	(0.69)	(0.79)
$\overline{Y}$	15.166	15.787	33.666
N	1176	1181	1151
R^2	0.267	0.113	0.199
	Panel C: Married-Mother Families Only		
Married-Mother Sample	0.016	1.03**	0.70
	(0.42)	(0.42)	(0.56)
$\overline{Y}$	18.145	17.659	37.476
N	3047	3024	2984
R^2	0.250	0.082	0.199

Note: Standard errors in parentheses and are clustered at the child-level; $*p < 0.10, **p < 0.05, ***p < 0.010$. Each cell represents a separate regression with full set of control variables, including child's gender, race, age, age square; mother's race, age, age square, marital status, educational attainment, whether mother is unmarried at child birth; number of children and age of the youngest child in the household; region, state, and year dummy; state policy and characteristics including per-capital real GDP, unemployment rate, real state minimum wage, combined state and federal current per pupil real spending on K-12 public education, combined monthly maximum AFDC/TANF and Food Stamps benefits for 3-person family. The full results including coefficients for all variables are in the online appendix or available upon request.

Table 2.8: Impact of EITC on Maternal Time Input

(Original CDS - Full Sample, Child FE)

	Engaged Time	Accessible Time	Total Time
	Panel A: Full CDS Families		
Full Sample	0.63	0.91	1.87**
	(0.50)	(0.58)	(0.76)
$\overline{Y}$	17.316	17.133	36.416
N	4223	4205	4135
R^2	0.247	0.064	0.194
	Panel B: Single-Mother Families Only		
Single-Mother Sample	-0.70	0.086	0.44
	(1.14)	(1.44)	(1.77)
$\overline{Y}$	15.166	15.787	33.666
N	1176	1181	1151
R^2	0.225	0.126	0.285
	Panel C: Married-Mother Families Only		
Married-Mother Sample	1.30**	0.83	2.13**
	(0.63)	(0.69)	(0.93)
$\overline{Y}$	18.145	17.659	37.476
N	3047	3024	2984
R^2	0.267	0.071	0.208

Note: Standard errors in parentheses and are clustered at the child-level; $*p < 0.10, **p < 0.05, ***p < 0.010$. Each cell represents a separate regression with full set of control variables, including child's gender, race, age, age square; mother's race, age, age square, marital status, educational attainment, whether mother is unmarried at child birth; number of children and age of the youngest child in the household; region, state, and year dummy; state policy and characteristics including per-capital real GDP, unemployment rate, real state minimum wage, combined state and federal current per pupil real spending on K-12 public education, combined monthly maximum AFDC/TANF and Food Stamps benefits for 3-person family. The full results including coefficients for all variables are in the online appendix or available upon request.

Table 2.9: Impact of EITC on Non-parental Child Care

(Augmented CDS - Full Sample, Pooled OLS with State FE)

	Any Child Care	CC by Relative	Formal CC	Weekly Hours
		Panel A: Full CDS Families		
Full Sample	0.027**	0.0022	0.018**	2.34***
	(0.012)	(0.010)	(0.0085)	(0.77)
$\overline{Y}$	0.368	0.134	0.161	16.563
N	17390	17380	17390	17369
R^2	0.161	0.058	0.111	0.153
		Panel B: Single-Mother Families Only		
Single-Mother Sample	0.037	0.026	-0.014	2.28*
	(0.025)	(0.018)	(0.014)	(1.23)
$\overline{Y}$	0.376	0.164	0.157	16.603
N	4458	4458	4458	4453
R^2	0.186	0.111	0.134	0.188
		Panel C: Married-Mother Families Only		
Married-Mother Sample	0.025*	-0.0040	0.032***	2.46**
	(0.013)	(0.012)	(0.0090)	(0.97)
$\overline{Y}$	0.366	0.124	0.163	16.550
N	12932	12922	12932	12916
R^2	0.175	0.065	0.120	0.167

Note: Standard errors in parentheses and are clustered at the state-level; * p<0.1, ** p<0.05, *** p<0.01. Each cell represents a separate regression with full set of control variables, including child's gender, race, age, age square; mother's race, age, age square, marital status, educational attainment, whether mother is unmarried at child birth; number of children and age of the youngest child in the household; year dummy; state policy and characteristics including per-capital real GDP, unemployment rate, real state minimum wage, combined state and federal current per pupil real spending on K-12 public education, combined monthly maximum AFDC/TANF and Food Stamps benefits for 3-person family. The full results including coefficients for all variables are available upon request.

Table 2.10: Impact of EITC on Non-parental Child Care

(Augmented CDS - Full Sample, Child FE)

	Any Child Care	CC by Relative	Formal CC	Weekly Hours
	Panel A: Full CDS Families			
Full Sample	0.043***	0.015*	0.042***	3.65***
	(0.012)	(0.0084)	(0.0094)	(0.76)
$\overline{Y}$	0.368	0.134	0.161	16.563
N	17390	17380	17390	17369
R^2	0.141	0.024	0.074	0.157
	Panel B: Single-Mother Families Only			
Single-Mother Sample	0.058**	0.037**	0.029	4.10**
	(0.023)	(0.017)	(0.019)	(1.75)
$\overline{Y}$	0.376	0.164	0.157	16.603
N	4458	4458	4458	4453
R^2	0.104	0.053	0.067	0.138
	Panel C: Married-Mother Families Only			
Married-Mother Sample	0.038***	0.0053	0.048***	3.44***
	(0.014)	(0.0095)	(0.011)	(0.83)
$\overline{Y}$	0.366	0.124	0.163	16.550
N	12932	12922	12932	12916
R^2	0.163	0.026	0.079	0.165

Note: Standard errors in parentheses and are clustered at the child-level; * p<0.1, ** p<0.05, *** p<0.01. Each cell represents a separate regression with full set of control variables, including child's gender, race, age, age square; mother's race, age, age square, marital status, educational attainment, whether mother is unmarried at child birth; number of children and age of the youngest child in the household; year dummy; state policy and characteristics including per-capital real GDP, unemployment rate, real state minimum wage, combined state and federal current per pupil real spending on K-12 public education, combined monthly maximum AFDC/TANF and Food Stamps benefits for 3-person family. The full results including coefficients for all variables are available upon request.

Table 2.11: Impact of EITC on Psychological Distress

(Original CDS - Full Sample, Pooled OLS with State FE)

	Kessler-6 Psychological Distress Scale
	Panel A: Full CDS Families
Full Sample	0.00023
	(0.11)
$\overline{Y}$	3.963
N	4334
R^2	0.081
	Panel B: Single-Mother Families Only
Single-Mother Sample	0.21
	(0.22)
$\overline{Y}$	4.857
N	1290
R^2	0.080
	Panel C: Married-Mother Families Only
Married-Mother Sample	-0.077
	(0.16)
$\overline{Y}$	3.583
N	3044
R^2	0.062

Note: Standard errors in parentheses and are clustered at the state-level; * p<0.1, ** p<0.05, *** p<0.01. Each cell represents a separate regression with full set of control variables, including child's gender, race, age, age square; mother's race, age, age square, marital status, educational attainment, whether mother is unmarried at child birth; number of children and age of the youngest child in the household; year dummy; state policy and characteristics including per-capital real GDP, unemployment rate, real state minimum wage, combined state and federal current per pupil real spending on K-12 public education, combined monthly maximum AFDC/TANF and Food Stamps benefits for 3-person family. The full results including coefficients for all variables are available upon request.

Table 2.12: Impact of EITC on Psychological Distress

(Original CDS - Full Sample, Child FE)

	Kessler-6 Psychological Distress Scale
	Panel A: Full CDS Families
Full Sample	0.16
	(0.13)
$\overline{Y}$	3.963
N	4334
R^2	0.021
	Panel B: Single-Mother Families Only
Single-Mother Sample	1.03***
	(0.38)
$\overline{Y}$	4.857
N	1290
R^2	0.104
	Panel C: Married-Mother Families Only
Married-Mother Sample	-0.16
	(0.15)
$\overline{Y}$	3.583
N	3044
R^2	0.030

Note: Standard errors in parentheses and are clustered at the state-level; * p<0.1, ** p<0.05, *** p<0.01. Each cell represents a separate regression with full set of control variables, including child's gender, race, age, age square; mother's race, age, age square, marital status, educational attainment, whether mother is unmarried at child birth; number of children and age of the youngest child in the household; year dummy; state policy and characteristics including per-capital real GDP, unemployment rate, real state minimum wage, combined state and federal current per pupil real spending on K-12 public education, combined monthly maximum AFDC/TANF and Food Stamps benefits for 3-person family. The full results including coefficients for all variables are available upon request.

2.13 Appendix

2.13.1 Details on Data and Variables

PSID-CDS

Table 2.13: Data Availability in the PSID-CDS

Variables	Source	Available Year
Child Human Capital Outcomes	CDS	1997, 2002, 2007
Maternal Time Investment	CDS-TD	1997, 2002, 2007
Non-parental Child Care	CDS	1997, 2002, 2007
Maternal Labor Market Outcomes	PSID	1985-1997, biennial afterwards
Household Characteristics	PSID	1985-1997, biennial afterwards

Notes: (i) Non-parental child care information is obtained based on retrospective questions on all arrangements used from birth until kindergarten enrollment and questions on the arrangement used at the time of survey. (ii) Child human capital outcomes are only available for children in older year, i.e. at least age 3 or 6 depending on specific measure.

In CDS-I, among the 2,803 children who are eligible to participate in any of the WJ-R subtests, 2,223 children attempted the assessment and 2,190 children completed all appropriate WJ-R subtests [30]. The detailed counts for the number of children who

[30]Potential reasons for a child not to complete the testing include: Caregiver terminated the test; Child would not respond; Major interruption; Child did not understand the WJ-R; A language problem; An emotional condition; A physical condition; Child was tired; or Other reason. In subsequent waves, a couple of factors lead to change in the number of children eligible for testing based on Duffy and Sastry (2014). Firstly, younger children who are too young to participate in

Table 2.14: Measurements of Cognitive Skills

WJ-R Subtest	Description	Age
Letter-Word Identification	Assesses symbolic learning and reading identification skills	≥ 3
Passage Comprehension	Assesses reading comprehension and vocabulary	≥ 6
Applied Problems	Assesses mathematics reasoning, achievement, and knowledge	≥ 3

Notes: i. Passage Comprehension is also available for Primary Care Giver in CDS 1997. 2. Another math-related test, Calculations, is only available in CDS 1997 but not accessed in either CDS 2002 or CDS 2007.

are eligible, who have started and who have completed WJ-R subtests are listed in Table 2.18.

CDS-I become eligible for the tests in CDS-II, i.e. those younger than age 3 or age 6 for different tests. Secondly, older children who reach age 18 would leave the CDS sample. Thirdly, some children have been lost to follow-up due to family's attrition from the Core PSID or from the CDS.

Table 2.15: Measurements of Non-cognitive Skills

Measure	Description	Age
Externalizing BPI	external or aggressive behaviors	≥ 3
Internalizing BPI	internal or withdrawn behaviors	≥ 3
Positive Behaviors Scale	emotional or social competence	≥ 3

Table 2.16: An Illustration Example of Child Care History in CDS-I

Age in 1997	t=0	t=1	t=2	t=3	t=4	t=5
12	1985	**1986**	1987	1988	1989	**1990**
11	**1986**	1987	1988	1989	**1990**	1991
10	1987	1988	1989	**1990**	1991	1992
9	1988	1989	**1990**	1991	1992	**1993**
8	1989	**1990**	1991	1992	**1993**	1994
7	**1990**	1991	1992	**1993**	1994	1995
6	1991	1992	**1993**	1994	1995	1996
5	1992	**1993**	1994	1995	1996	1997
4	**1993**	1994	1995	1996	1997	1998
3	1994	1995	1996	1997	1998	1999
2	1995	1996	1997	1998	1999	2000
1	1996	1997	1998	1999	2000	2001
0	1997	1998	1999	2000	2001	2002

Table 2.17: Cohorts of Children in the PSID-CDS

Year of Birth	Child's Age and Corresponding Calendar Year												
t=0	t=1	t=2	t=3	···	t=7	t=8	t=9	···	t=12	t=13	t=14	t=15	t=16
1985	1986	1987	1988	···	1992	1993	1994	···	**1997**	1998	1999	2000	2001
1986	1987	1988	1989	···	1993	1994	1995	···	1998	1999	2000	2001	**2002**
1987	1988	1989	1990	···	1994	1995	1996	···	1999	2000	2001	**2002**	2003
1988	1989	1990	1991	···	1995	1996	**1997**	···	2000	2001	**2002**	2003	2004
1989	1990	1991	1992	···	1996	**1997**	1998	···	2001	**2002**	2003	2004	2005
1990	1991	1992	1993	···	**1997**	1998	1999	···	**2002**	2003	2004	2005	2006
1991	1992	1993	1994	···	1998	1999	2000	···	2003	2004	2005	2006	**2007**
1992	1993	1994	1995	···	1999	2000	2001	···	2004	2005	2006	**2007**	2008
1993	1994	1995	1996	···	2000	2001	**2002**	···	2005	2006	**2007**	2008	2009
1994	1995	1996	**1997**	···	2001	**2002**	2003	···	2006	**2007**	2008	2009	2010
1995	1996	**1997**	1998	···	**2002**	2003	2004	···	**2007**	2008	2009	2010	2011
1996	**1997**	1998	1999	···	2003	2004	2005	···	2008	2009	2010	2011	2012
1997	1998	1999	2000	···	2004	2005	2006	···	2009	2010	2011	2012	2013

Table 2.18: Counts, by Wave, of Children Eligible for WJ-R Subtests

	CDS-I (3563 children)			CDS-II (2907 children)			CDS-III (1608 children)		
WJ-R Subtest	Eligible	Began Subtest	Completed Subtest	Eligible	Began Subtest	Completed Subtest	Eligible	Began Subtest	Completed Subtest
Letter-Word Identification	2,803	2,223	2,218	2,907	2,644	2,633	1,608	1,491	1,490
Passage Comprehension	1,877†	1,532	1,522	2,856	2,595	2,541	1,608	1,491	1,491
Applied Problems	2,803	2,223	2,209	2,907	2,644	2,625	1,608	1,491	1,485
Calculations	1,877†	1,532	1,517	N/A	N/A	N/A	N/A	N/A	N/A

Notes: † Approximate number of eligible children for Passage Comprehension and Calculation subtests based on age at 1997 Core PSID interview. This is reproduced from Duffy and Sastry (2014)(Table 3).

The Aggregate Test Score in Bastian and Michelmore (2018)

In a recent paper by Bastian and Michelmore (2018), they use the PSID and its CDS to investigate the long-run impact of the EITC on children's education and employment outcomes. In "Table: Effect of EITC Exposure on Intermediate Outcomes (OLS Results)", they show that contemporaneous EITC generosity leads to an 0.126 increase of Standardized Test Scores at the statistical significance of 5%. However, this conclusion is questionable, because of the error in their construction of such "standardized test score".

However, such "standardized test score" measure may contain some issues, based on the description in "Appendix C: Description Description of Data and Variables" of their working paper[31].

Inconsistent and Misclassification in the Math score measure

1. `Q3BMA_SS` is the Broad Math Score 97, which contains Applied Problem and Calculation in CDS-97. This is the only available broad math variable in all three waves of CDS, because Calculation is only assessed in the first wave of CDS 1997. In other words, no consistent measure of such same variable is available in CDDS 02 and CDS 07. Moreover, this variable is objectively assessed.

2. `MATH02` and `MATH07` are Ability Self Concepts Math 02 and 07. These questions are answered by children subjectively.

[31] `https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2674603`

Hence, it is inappropriate and inconsistent to treat these three variables as if they are the same variable being measured in different waves. These two are essentially different variables.

Duplication in both Reading and Math score

1. The "Reading" (`Q24BRSS, Q3BRE_SS, Q34BRSS`) is already a a combination of both *Letter-Word* and *Passage Comprehension*. Therefore, including both such "Reading" score and "Passage Comprehension" cause duplication, i.e. *Passage Comprehension* is calculated twice in all three waves of CDS.

2. `Q3BMA_SS` is already a combination of both *Applied Problem* and *Calculation*. Therefore, including both such "Math" and "Applied Problem" cause duplication, i.e. *Applied Problem* is calculated twice in the first wave of CDS.

Hence, the authors are duplicating Passage Comprehension for twice for all CDS waves and duplicating the Applied Problem for once for CDS97.

2.13.2 Additional Results

Table 2.19: Understanding Sources of Variation

	Overall	Single-Mother	Married-Mother
Two Children or More	0.80***	0.79***	0.80***
	(0.0053)	(0.011)	(0.0051)
Year 2002	0.031***	0.0045	0.041***
	(0.0043)	(0.0098)	(0.0053)
Year 2007	0.056***	0.046***	0.058***
	(0.0075)	(0.015)	(0.0097)
Constant	2.43***	2.48***	2.43***
	(0.0088)	(0.018)	(0.0093)
Observations	7163	2092	4814
R^2 within Model	0.899	0.881	0.907
R^2 between Model	0.664	0.649	0.692
R^2 overall Model	0.721	0.679	0.750

Note: Standard errors in parentheses; $*p < 0.10, **p < 0.05, ***p < 0.010$.

Table 2.20: Impact of EITC on Child Developmental Outcomes

(Excluding High-Income Households (> $150k), Pooled OLS with State FE)

	Cognitive Skill				**Non-cognitive Skill**		
	Letter Word	Passage Comprehension	Broad Reading	Applied Problem	Externalizing Problem	Internalizing Problem	Positive Behavior
				Panel A: Full CDS Families			
EITC Generosity	-0.021	-0.020	-0.026	0.017	-0.050	-0.072**	0.022
	(0.040)	(0.027)	(0.032)	(0.032)	(0.034)	(0.029)	(0.030)
$\overline{Y}$	-0.003	0.018	0.004	0.039	-0.027	-0.018	0.009
N	4072	3549	3547	3602	4228	4236	4275
R^2	0.214	0.256	0.269	0.272	0.066	0.062	0.051
				Panel B: Single - Mother Families Only			
EITC Generosity	-0.083	-0.088*	-0.10**	-0.049	-0.058	-0.13**	0.0069
	(0.050)	(0.044)	(0.050)	(0.052)	(0.053)	(0.050)	(0.050)
$\overline{Y}$	-0.316	-0.331	-0.351	-0.273	0.143	0.044	-0.021
N	1308	1157	1156	1166	1362	1364	1379
R^2	0.175	0.221	0.228	0.221	0.076	0.116	0.084
				Panel C: Married - Mother Families Only			
EITC Generosity	0.0028	0.0021	0.0026	0.051	-0.062	-0.043	0.036
	(0.043)	(0.036)	(0.037)	(0.041)	(0.039)	(0.036)	(0.039)
$\overline{Y}$	0.146	0.187	0.175	0.188	-0.107	-0.047	0.023
N	2764	2392	2391	2436	2866	2872	2896
R^2	0.189	0.225	0.237	0.257	0.059	0.059	0.052

Note: Standard errors in parentheses and are clustered at the state-level; * p<0.1, ** p<0.05, *** p<0.01. Each cell represents a separate regression with full set of control variables, including child's gender, race, age, age square; mother's race, age, age square, marital status, educational attainment, whether mother is unmarried at child birth; number of children and age of the youngest child in the household; region, state, and year dummy; state policy and characteristics including per-capital real GDP, unemployment rate, real state minimum wage, combined state and federal current per pupil real spending on K-12 public education, combined monthly maximum AFDC/TANF and Food Stamps benefits for 3-person family. The full results including coefficients for all variables are available upon request.

Table 2.21: Impact of EITC on Child Developmental Outcomes

(Excluding High-Income Households (> $100k), Pooled OLS with State FE)

	Cognitive Skill				**Non-cognitive Skill**		
	Letter Word	Passage Comprehension	Broad Reading	Applied Problem	Externalizing Problem	Internalizing Problem	Positive Behavior
				Panel A: Full CDS Families			
EITC Generosity	-0.030	-0.024	-0.032	0.020	-0.054	-0.071**	0.0080
	(0.046)	(0.030)	(0.038)	(0.036)	(0.033)	(0.031)	(0.032)
$\overline{Y}$	-0.057	-0.036	-0.058	-0.029	0.002	-0.002	0.005
N	3657	3162	3161	3213	3788	3796	3834
R^2	0.206	0.255	0.263	0.250	0.065	0.066	0.053
				Panel B: Single - Mother Families Only			
EITC Generosity	-0.082	-0.10**	-0.11**	-0.055	-0.057	-0.12**	0.0027
	(0.050)	(0.042)	(0.050)	(0.051)	(0.053)	(0.052)	(0.051)
$\overline{Y}$	-0.322	-0.337	-0.358	-0.276	0.144	0.041	-0.018
N	1294	1143	1142	1152	1348	1350	1365
R^2	0.176	0.222	0.230	0.220	0.076	0.116	0.085
				Panel C: Married - Mother Families Only			
EITC Generosity	-0.0033	0.010	0.0060	0.069	-0.075**	-0.041	0.020
	(0.055)	(0.044)	(0.048)	(0.048)	(0.037)	(0.038)	(0.042)
$\overline{Y}$	0.088	0.134	0.112	0.109	-0.076	-0.027	0.018
N	2363	2019	2019	2061	2440	2446	2469
R^2	0.187	0.234	0.240	0.241	0.063	0.068	0.057

Note: Standard errors in parentheses and are clustered at the state-level; * p<0.1, ** p<0.05, *** p<0.01. Each cell represents a separate regression with full set of control variables, including child's gender, race, age, age square; mother's race, age, age square, marital status, educational attainment, whether mother is unmarried at child birth; number of children and age of the youngest child in the household; region, state, and year dummy; state policy and characteristics including per-capital real GDP, unemployment rate, real state minimum wage, combined state and federal current per pupil real spending on K-12 public education, combined monthly maximum AFDC/TANF and Food Stamps benefits for 3-person family. The full results including coefficients for all variables are available upon request.

Table 2.22: Impact of EITC on Child Developmental Outcomes

(Excluding High-Income Households (> $85k), Pooled OLS with State FE)

	Cognitive Skill				**Non-cognitive Skill**		
	Letter Word	Passage Comprehension	Broad Reading	Applied Problem	Externalizing Problem	Internalizing Problem	Positive Behavior
				Panel A: Full CDS Families			
EITC Generosity	-0.035	-0.023	-0.034	0.018	-0.044	-0.075**	-0.0024
	(0.050)	(0.034)	(0.042)	(0.036)	(0.034)	(0.032)	(0.032)
$\overline{Y}$	-0.088	-0.072	-0.094	-0.063	0.017	0.004	0.004
N	3416	2938	2937	2985	3542	3548	3585
R^2	0.204	0.248	0.259	0.240	0.065	0.069	0.055
				Panel B: Single - Mother Families Only			
EITC Generosity	-0.080	-0.097**	-0.11**	-0.049	-0.059	-0.12**	0.00055
	(0.052)	(0.044)	(0.052)	(0.051)	(0.053)	(0.053)	(0.051)
$\overline{Y}$	-0.329	-0.340	-0.364	-0.279	0.146	0.044	-0.020
N	1282	1131	1130	1140	1336	1338	1353
R^2	0.177	0.225	0.231	0.221	0.075	0.116	0.084
				Panel C: Married - Mother Families Only			
EITC Generosity	-0.013	0.0028	-0.0011	0.066	-0.058	-0.041	0.0069
	(0.061)	(0.048)	(0.054)	(0.051)	(0.037)	(0.039)	(0.043)
$\overline{Y}$	0.057	0.095	0.075	0.071	-0.061	-0.021	0.018
N	2134	1807	1807	1845	2206	2210	2232
R^2	0.188	0.234	0.242	0.231	0.064	0.070	0.061

Note: Standard errors in parentheses and are clustered at the state-level; * p<0.1, ** p<0.05, *** p<0.01. Each cell represents a separate regression with full set of control variables, including child's gender, race, age, age square; mother's race, age, age square, marital status, educational attainment, whether mother is unmarried at child birth; number of children and age of the youngest child in the household; region, state, and year dummy; state policy and characteristics including per-capital real GDP, unemployment rate, real state minimum wage, combined state and federal current per pupil real spending on K-12 public education, combined monthly maximum AFDC/TANF and Food Stamps benefits for 3-person family. The full results including coefficients for all variables are available upon request.

Table 2.23: Impact of EITC on Child Developmental Outcomes

(Only High-school Dropouts, Pooled OLS with State FE)

	Cognitive Skill				**Non-cognitive Skill**		
	Letter Word	Passage Comprehension	Broad Reading	Applied Problem	Externalizing Problem	Internalizing Problem	Positive Behavior
				Panel A: Full CDS Families			
EITC Generosity	-0.053	0.019	-0.032	0.032	0.032	-0.068	0.0065
	(0.10)	(0.074)	(0.096)	(0.063)	(0.074)	(0.085)	(0.093)
$\overline{Y}$	-0.473	-0.530	-0.545	-0.455	0.116	0.124	0.095
N	531	462	462	472	552	557	561
R^2	0.255	0.295	0.318	0.219	0.159	0.126	0.116
				Panel B: Single - Mother Families Only			
EITC Generosity	-0.24**	-0.17*	-0.24***	-0.084	-0.030	-0.15	0.20*
	(0.091)	(0.092)	(0.087)	(0.073)	(0.12)	(0.11)	(0.10)
$\overline{Y}$	-0.710	-0.707	-0.764	-0.551	0.286	0.201	-0.002
N	239	208	208	212	249	252	255
R^2	0.298	0.435	0.409	0.381	0.185	0.231	0.144
				Panel C: Married - Mother Families Only			
EITC Generosity	-0.031	0.093	0.023	0.11	-0.065	-0.10	0.013
	(0.14)	(0.072)	(0.11)	(0.092)	(0.095)	(0.13)	(0.12)
$\overline{Y}$	-0.280	-0.385	-0.365	-0.377	-0.024	0.061	0.176
N	292	254	254	260	303	305	306
R^2	0.291	0.342	0.355	0.217	0.281	0.197	0.215

Note: Standard errors in parentheses and are clustered at the state-level; * p<0.1, ** p<0.05, *** p<0.01. Each cell represents a separate regression with full set of control variables, including child's gender, race, age, age square; mother's race, age, age square, marital status, educational attainment, whether mother is unmarried at child birth; number of children and age of the youngest child in the household; region, state, and year dummy; state policy and characteristics including per-capital real GDP, unemployment rate, real state minimum wage, combined state and federal current per pupil real spending on K-12 public education, combined monthly maximum AFDC/TANF and Food Stamps benefits for 3-person family. The full results including coefficients for all variables are available upon request.

Table 2.24: Impact of EITC on Child Developmental Outcomes

(Only High-school Graduates and Above, Pooled OLS with State FE)

	Cognitive Skill				**Non-cognitive Skill**		
	Letter Word	Passage Comprehension	Broad Reading	Applied Problem	Externalizing Problem	Internalizing Problem	Positive Behavior
				Panel A: Full CDS Families			
EITC Generosity	-0.012	-0.037	-0.027	0.0043	-0.064**	-0.076***	0.027
	(0.033)	(0.028)	(0.028)	(0.030)	(0.030)	(0.027)	(0.028)
$\overline{Y}$	0.107	0.131	0.127	0.163	-0.067	-0.046	0.000
N	3765	3300	3298	3343	3913	3917	3952
R^2	0.204	0.238	0.254	0.278	0.063	0.061	0.050
				Panel B: Single - Mother Families Only			
EITC Generosity	-0.061	-0.059	-0.073	-0.038	-0.076	-0.14**	-0.015
	(0.044)	(0.046)	(0.046)	(0.050)	(0.053)	(0.055)	(0.045)
$\overline{Y}$	-0.227	-0.247	-0.259	-0.208	0.109	0.007	-0.024
N	1072	951	950	956	1116	1115	1127
R^2	0.148	0.193	0.201	0.218	0.073	0.117	0.103
				Panel C: Married - Mother Families Only			
EITC Generosity	0.0019	-0.037	-0.016	0.023	-0.062*	-0.042	0.040
	(0.038)	(0.032)	(0.034)	(0.037)	(0.033)	(0.031)	(0.036)
$\overline{Y}$	0.239	0.285	0.283	0.312	-0.137	-0.068	0.010
N	2693	2349	2348	2387	2797	2802	2825
R^2	0.186	0.204	0.225	0.253	0.055	0.059	0.047

Note: Standard errors in parentheses and are clustered at the state-level; * p<0.1, ** p<0.05, *** p<0.01. Each cell represents a separate regression with full set of control variables, including child's gender, race, age, age square; mother's race, age, age square, marital status, educational attainment, whether mother is unmarried at child birth; number of children and age of the youngest child in the household; region, state, and year dummy; state policy and characteristics including per-capital real GDP, unemployment rate, real state minimum wage, combined state and federal current per pupil real spending on K-12 public education, combined monthly maximum AFDC/TANF and Food Stamps benefits for 3-person family. The full results including coefficients for all variables are available upon request.

Table 2.25: Impact of EITC on Child Developmental Outcomes

(Excluding High-Income Households (> $150k), Child FE)

	Cognitive Skill				**Non-cognitive Skill**		
	Letter Word	Passage Comprehension	Broad Reading	Applied Problem	Externalizing Problem	Internalizing Problem	Positive Behavior
				Panel A: Full CDS Families			
Full Sample	-0.055	-0.048	-0.035	0.0033	-0.065*	-0.090**	0.077*
	(0.035)	(0.043)	(0.035)	(0.036)	(0.035)	(0.037)	(0.043)
$\overline{Y}$	-0.003	0.018	0.004	0.039	-0.027	-0.018	0.009
N	4072	3549	3547	3602	4228	4236	4275
R^2	0.048	0.057	0.050	0.117	0.022	0.043	0.010
				Panel B: Single - Mother Families Only			
Single-Mother Sample	-0.11*	-0.13	-0.13**	-0.0036	-0.23***	-0.20**	0.089
	(0.064)	(0.083)	(0.064)	(0.086)	(0.084)	(0.086)	(0.11)
$\overline{Y}$	-0.316	-0.331	-0.351	-0.273	0.143	0.044	-0.021
N	1308	1157	1156	1166	1362	1364	1379
R^2	0.068	0.119	0.113	0.167	0.039	0.080	0.041
				Panel C: Married - Mother Families Only			
Married-Mother Sample	-0.037	-0.011	-0.0046	0.019	-0.019	-0.043	0.12**
	(0.048)	(0.060)	(0.049)	(0.049)	(0.044)	(0.045)	(0.048)
$\overline{Y}$	0.146	0.187	0.175	0.188	-0.107	-0.047	0.023
N	2764	2392	2391	2436	2866	2872	2896
R^2	0.064	0.058	0.062	0.125	0.033	0.049	0.017

Note: Standard errors in parentheses and are clustered at the state-level; * p<0.1, ** p<0.05, *** p<0.01. Each cell represents a separate regression with full set of control variables, including child's gender, race, age, age square; mother's race, age, age square, marital status, educational attainment, whether mother is unmarried at child birth; number of children and age of the youngest child in the household; region, state, and year dummy; state policy and characteristics including per-capital real GDP, unemployment rate, real state minimum wage, combined state and federal current per pupil real spending on K-12 public education, combined monthly maximum AFDC/TANF and Food Stamps benefits for 3-person family. The full results including coefficients for all variables are available upon request.

Table 2.26: Impact of EITC on Child Developmental Outcomes

(Excluding High-Income Households (> $100k), Child FE)

	Cognitive Skill				**Non-cognitive Skill**		
	Letter Word	Passage Comprehension	Broad Reading	Applied Problem	Externalizing Problem	Internalizing Problem	Positive Behavior
				Panel A: Full CDS Families			
Full Sample	-0.054	-0.050	-0.043	0.00017	-0.086**	-0.091**	0.079
	(0.038)	(0.046)	(0.038)	(0.039)	(0.039)	(0.040)	(0.049)
$\overline{Y}$	-0.057	-0.036	-0.058	-0.029	0.002	-0.002	0.005
N	3657	3162	3161	3213	3788	3796	3834
R^2	0.051	0.068	0.063	0.113	0.021	0.046	0.011
				Panel B: Single - Mother Families Only			
Single-Mother Sample	-0.10	-0.11	-0.11*	0.00083	-0.24***	-0.22**	0.078
	(0.065)	(0.083)	(0.064)	(0.086)	(0.085)	(0.087)	(0.12)
$\overline{Y}$	-0.322	-0.337	-0.358	-0.276	0.144	0.041	-0.018
N	1294	1143	1142	1152	1348	1350	1365
R^2	0.068	0.118	0.115	0.165	0.041	0.079	0.038
				Panel C: Married - Mother Families Only			
Married-Mother Sample	-0.043	-0.023	-0.027	0.0033	-0.030	-0.029	0.14**
	(0.054)	(0.066)	(0.056)	(0.054)	(0.051)	(0.050)	(0.057)
$\overline{Y}$	0.088	0.134	0.112	0.109	-0.076	-0.027	0.018
N	2363	2019	2019	2061	2440	2446	2469
R^2	0.069	0.070	0.072	0.111	0.031	0.054	0.022

Note: Standard errors in parentheses and are clustered at the state-level; * p<0.1, ** p<0.05, *** p<0.01. Each cell represents a separate regression with full set of control variables, including child's gender, race, age, age square; mother's race, age, age square, marital status, educational attainment, whether mother is unmarried at child birth; number of children and age of the youngest child in the household; region, state, and year dummy; state policy and characteristics including per-capital real GDP, unemployment rate, real state minimum wage, combined state and federal current per pupil real spending on K-12 public education, combined monthly maximum AFDC/TANF and Food Stamps benefits for 3-person family. The full results including coefficients for all variables are available upon request.

Table 2.27: Impact of EITC on Child Developmental Outcomes

(Excluding High-Income Households (> $85k), Child FE)

	Cognitive Skill				**Non-cognitive Skill**		
	Letter Word	Passage Comprehension	Broad Reading	Applied Problem	Externalizing Problem	Internalizing Problem	Positive Behavior
				Panel A: Full CDS Families			
Full Sample	-0.064*	-0.066	-0.062	0.011	-0.10**	-0.11**	0.065
	(0.039)	(0.050)	(0.041)	(0.042)	(0.042)	(0.043)	(0.052)
$\overline{Y}$	-0.088	-0.072	-0.094	-0.063	0.017	0.004	0.004
N	3416	2938	2937	2985	3542	3548	3585
R^2	0.054	0.076	0.067	0.111	0.021	0.045	0.012
				Panel B: Single - Mother Families Only			
Single-Mother Sample	-0.10	-0.11	-0.12*	-0.00084	-0.25***	-0.22**	0.084
	(0.065)	(0.083)	(0.064)	(0.086)	(0.085)	(0.087)	(0.12)
$\overline{Y}$	-0.329	-0.340	-0.364	-0.279	0.146	0.044	-0.020
N	1282	1131	1130	1140	1336	1338	1353
R^2	0.068	0.120	0.117	0.165	0.043	0.078	0.039
				Panel C: Married - Mother Families Only			
Married-Mother Sample	-0.061	-0.041	-0.046	0.0097	-0.048	-0.035	0.11*
	(0.058)	(0.075)	(0.064)	(0.060)	(0.054)	(0.056)	(0.063)
$\overline{Y}$	0.057	0.095	0.075	0.071	-0.061	-0.021	0.018
N	2134	1807	1807	1845	2206	2210	2232
R^2	0.071	0.093	0.080	0.108	0.031	0.057	0.023

Note: Standard errors in parentheses and are clustered at the state-level; * p<0.1, ** p<0.05, *** p<0.01. Each cell represents a separate regression with full set of control variables, including child's gender, race, age, age square; mother's race, age, age square, marital status, educational attainment, whether mother is unmarried at child birth; number of children and age of the youngest child in the household; region, state, and year dummy; state policy and characteristics including per-capital real GDP, unemployment rate, real state minimum wage, combined state and federal current per pupil real spending on K-12 public education, combined monthly maximum AFDC/TANF and Food Stamps benefits for 3-person family. The full results including coefficients for all variables are available upon request.

Table 2.28: Impact of EITC on Child Developmental Outcomes

(Only High-school Dropouts, Child FE)

	Cognitive Skill				**Non-cognitive Skill**		
	Letter Word	Passage Comprehension	Broad Reading	Applied Problem	Externalizing Problem	Internalizing Problem	Positive Behavior
				Panel A: Full CDS Families			
Full Sample	0.037	-0.24**	-0.11	0.015	-0.14	-0.028	-0.0028
	(0.12)	(0.11)	(0.11)	(0.100)	(0.13)	(0.14)	(0.15)
$\overline{Y}$	-0.473	-0.530	-0.545	-0.455	0.116	0.124	0.095
N	531	462	462	472	552	557	561
R^2	0.123	0.254	0.261	0.254	0.080	0.097	0.099
				Panel B: Single - Mother Families Only			
Single-Mother Sample	-0.23	-0.64***	-0.51***	-0.22	-0.41*	-0.31	0.00012
	(0.15)	(0.16)	(0.11)	(0.14)	(0.22)	(0.25)	(0.32)
$\overline{Y}$	-0.710	-0.707	-0.764	-0.551	0.286	0.201	-0.002
N	239	208	208	212	249	252	255
R^2	0.196	0.393	0.326	0.530	0.159	0.137	0.147
				Panel C: Married - Mother Families Only			
Married-Mother Sample	0.020	-0.090	-0.040	0.17	0.093	0.36**	0.15
	(0.23)	(0.15)	(0.16)	(0.15)	(0.18)	(0.18)	(0.15)
$\overline{Y}$	-0.280	-0.385	-0.365	-0.377	-0.024	0.061	0.176
N	292	254	254	260	303	305	306
R^2	0.221	0.425	0.452	0.271	0.131	0.179	0.188

Note: Standard errors in parentheses and are clustered at the state-level; * p<0.1, ** p<0.05, *** p<0.01. Each cell represents a separate regression with full set of control variables, including child's gender, race, age, age square; mother's race, age, age square, marital status, educational attainment, whether mother is unmarried at child birth; number of children and age of the youngest child in the household; region, state, and year dummy; state policy and characteristics including per-capital real GDP, unemployment rate, real state minimum wage, combined state and federal current per pupil real spending on K-12 public education, combined monthly maximum AFDC/TANF and Food Stamps benefits for 3-person family. The full results including coefficients for all variables are available upon request.

Table 2.29: Impact of EITC on Child Developmental Outcomes

(Only High-school Graduates and Above, Child FE)

	Cognitive Skill				**Non-cognitive Skill**		
	Letter Word	Passage Comprehension	Broad Reading	Applied Problem	Externalizing Problem	Internalizing Problem	Positive Behavior
				Panel A: Full CDS Families			
Full Sample	-0.051	-0.072	-0.042	-0.0080	-0.055	-0.10***	0.086**
	(0.036)	(0.044)	(0.035)	(0.035)	(0.036)	(0.037)	(0.041)
$\overline{Y}$	0.107	0.131	0.127	0.163	-0.067	-0.046	0.000
N	3765	3300	3298	3343	3913	3917	3952
R^2	0.052	0.049	0.045	0.113	0.022	0.044	0.013
				Panel B: Single - Mother Families Only			
Single-Mother Sample	-0.077	-0.044	-0.065	0.057	-0.18*	-0.16	0.12
	(0.067)	(0.093)	(0.074)	(0.099)	(0.096)	(0.098)	(0.12)
$\overline{Y}$	-0.227	-0.247	-0.259	-0.208	0.109	0.007	-0.024
N	1072	951	950	956	1116	1115	1127
R^2	0.096	0.106	0.110	0.152	0.050	0.108	0.068
				Panel C: Married - Mother Families Only			
Married-Mother Sample	-0.035	-0.069	-0.031	-0.0029	-0.041	-0.097**	0.10**
	(0.047)	(0.059)	(0.047)	(0.046)	(0.043)	(0.044)	(0.043)
$\overline{Y}$	0.239	0.285	0.283	0.312	-0.137	-0.068	0.010
N	2693	2349	2348	2387	2797	2802	2825
R^2	0.069	0.057	0.068	0.119	0.029	0.040	0.023

Note: Standard errors in parentheses and are clustered at the state-level; * p<0.1, ** p<0.05, *** p<0.01. Each cell represents a separate regression with full set of control variables, including child's gender, race, age, age square; mother's race, age, age square, marital status, educational attainment, whether mother is unmarried at child birth; number of children and age of the youngest child in the household; region, state, and year dummy; state policy and characteristics including per-capital real GDP, unemployment rate, real state minimum wage, combined state and federal current per pupil real spending on K-12 public education, combined monthly maximum AFDC/TANF and Food Stamps benefits for 3-person family. The full results including coefficients for all variables are available upon request.

Table 2.30: Impact of EITC on Child Developmental Outcomes

Seemingly Unrelated Regression (SUR)

	Full Sample	**Single Mother**	**Married Mother**
Broad Reading	-0.030	-0.10**	-0.0096
	(0.026)	(0.047)	(0.032)
BPI: Internalizing	-0.087***	-0.15***	-0.055
	(0.030)	(0.053)	(0.036)
BPI: Externalizing	-0.084***	-0.090*	-0.094***
	(0.029)	(0.053)	(0.035)
$\overline{Y}$	0.047	-0.347	0.221
N	3707	1138	2569

Note: Standard errors in parentheses; * p<0.1, ** p<0.05, *** p<0.01. Each cell represents a separate regression with full set of control variables, including child's gender, race, age, age square; mother's race, age, age square, marital status, educational attainment, whether mother is unmarried at child birth; number of children and age of the youngest child in the household; region, state, and year dummy; state policy and characteristics including per-capital real GDP, unemployment rate, real state minimum wage, combined state and federal current per pupil real spending on K-12 public education, combined monthly maximum AFDC/TANF and Food Stamps benefits for 3-person family. The full results including coefficients for all variables are available upon request.

Chapter 3

Paternal Migration, Investment in Children, and Children's Non-cognitive Development: Evidence from Rural China

This chapter is joint work with Xi Yang.[1]

[1]We are grateful to our advisor Robert Moffitt for his encouragement and guidance. We also thank Sarah Baird, David Bishai, Barry Chiswick, Shahe Emran, James Foster, Edward Glaeser, Yingyao Hu, Remi Jedwab, Matthew Kahn, Dave Marcotte, Erdal Tekin, Michael Nieswiadomy, Bryan Stuart, and Junfu Zhang, as well as seminar participants at the Economics Department seminar of University of North Texas, Chinese Economists Society 2019 North America Conference, 2018 GW Student Research Conference, 2018 Association for Public Policy Analysis and Management DC Regional Student Conference, and the 4th Biennial Conference of China Development Studies for helpful comments and suggestions. All errors are our own. The Gansu Survey of Children and Families is supported by a grant from the United Kingdom Economic and Social Research Council and the Department for International Development (ESRC RES-167-25-0250). Earlier support for data collection came from the Spencer Foundation Small and Major Grants Programs, World Bank, and National Institutes for Health grants 1R01TW005930-01 and 5R01TW005930-02. We

CHAPTER 3. PATERNAL MIGRATION, INVESTMENT IN CHILDREN, AND CHILDREN'S NON-COGNITIVE DEVELOPMENT: EVIDENCE FROM RURAL CHINA

Abstract: Many children worldwide are left behind by parents who are migrating for work. While previous literature has studied the effect of parental migration on children's educational outcomes and cognitive achievements, this study focuses on how paternal migration affects children's non-cognitive development. We use longitudinal data of children in rural China and adopt labor market conditions in destination provinces as instrumental variables for parental endogenous migration choice. We find that father's migration has a significant negative effect on children's non-cognitive development, particularly for boys. We test and discuss several mechanisms including parental financial inputs, parental time inputs, children's own time input, household bargaining, and parenting style.

Keywords: Left-behind Children, Parental Migration, Parental Input, Non-cognitive Development, Rural China

JEL Classification: J12, J13, J24, J61, R23

3.1 Introduction

International and internal migrants together account for about 1 billion people worldwide (United Nations, 2013). Migrants often leave their children behind because of rigid migration policies, high cost of migration, or uncertain living conditions in the destination region. For example, in China, about 61 million children are left

appreciate Emily Hannum and Albert Park for generously providing data.

behind in rural areas by their parents who have migrated in search of jobs in urban cities, with an increasing trend of one-parent migration (usually the father).[2] This "left-behind children" phenomenon has received growing attention in the literature. So far, substantial research has studied the effect of parental migration on left-behind children's cognitive, educational, and health outcomes.[3] Yet, studies on the causal impacts of parental migration on left-behind children's non-cognitive development remain limited.[4] Moreover, we know even less about the underlying mechanisms through which parental migration affect children's development.

This paper is among the first to examine the impact of father's migration on left-behind children's non-cognitive outcomes by using the Gansu Survey of Children and Families (GSCF), a longitudinal study of children between 9 and 16 years old in rural China. The GSCF is one of the few datasets that measure children's non-cognitive skills in a developing country context, which allows us to expand the scope of existing studies by linking parental migration to children's non-cognitive development. More important, by taking advantage of the richness of the GSCF, this paper can evaluate the mechanisms behind parental migration in a more comprehensive way. Our results

[2]This number is estimated based on the 2010 population census by the All-China Women's Federation.

[3]These studies found that parental migration has significant effects on education outcomes (Acosta, 2011; Alcaraz et al., 2012; Amuedo-Dorantes and Pozo, 2010; Antman, 2011, 2012; Binci and Giannelli, 2018; Cortes, 2015; Cox-Edwards and Ureta, 2003; Giannelli and Mangiavacchi, 2010; Gibson and McKenzie, 2014; Kroeger and Anderson, 2014; Macours and Vakis, 2010; McKenzie and Rapoport, 2011; Nobles, 2011; Powers, 2011; Yang, 2008; Zhang et al., 2014) and health outcomes (De Brauw and Mu, 2011; Gibson et al., 2011b; Hildebrandt and McKenzie, 2005; Mu and De Brauw, 2015; Stillman et al., 2012). Also see Antman (2013) for a survey of the effects of migrant parents on the left-behind family.

[4]This is partially because of the lack of reliable data.

are not only important for understanding the consequences of parental migration, which is major phenomenon in developing countries, but also add to our knowledge about the formation of children's non-cognitive skills in general.

Non-cognitive skills can be broadly defined as social-emotional skills, which are usually referred to as the second dimension of individual heterogeneity in addition to cognitive skills in labor economics.[5] Understanding the impact of parental migration on children's non-cognitive skills is of particular interest given growing evidence that non-cognitive skills have a large impact on economic outcomes (Flossmann et al., 2006; Heckman and Rubinstein, 2001; Heckman et al., 2006; Kautz et al., 2014; Lindqvist and Vestman, 2011; Weinberger, 2014) and are more malleable in later childhood than cognitive skills (Akee et al., 2018; Almlund et al., 2011; Borghans et al., 2008).

Theoretically, in an intact rural family, father's migration can affect the non-cognitive development of children left behind through several mechanisms. Financial inputs and time inputs are two such mechanisms that have been heavily discussed in the literature. First, migration is usually motivated by and associated with higher income. Thus, migrant parents can send home remittances that are higher than the wages they could have earned at home. The additional financial resources can have a positive impact on children's non-cognitive development by reducing emotional and

[5]These skills have many names in the literature including soft skills, personality traits, non-cognitive abilities, character skills, and socio-emotional skills. See Almlund et al. (2011); Borghans et al. (2008); Humphries and Kosse (2017); Thiel and Thomsen (2013) for measurements and interpretations of non-cognitive skills across different disciplines such as psychologic, sociologic and economics.

behavioral problems caused by financial distress.[6]

Second, parental migration likely reduces parental time inputs - not only paternal time input directly, but also maternal time input indirectly[7], which may have a negative impact on children's non-cognitive development, especially considering that several studies have documented that parental time investments are vital for children's non-cognitive outcomes (Bono et al., 2016; Cunha and Heckman, 2008; Cunha et al., 2010; Fiorini and Keane, 2014; Hsin and Felfe, 2014).[8] Since the above two mechanisms work in opposite directions, the effect of parental migration on children's non-cognitive development is theoretically ambiguous and can only be determined empirically.

The tradeoff between financial and time inputs has been intensively investigated in the literature studying the effect of mother's labor supply and children's human capital outcomes (See Heckman and Mosso (2014) for a recent review). These papers argue that an increase in maternal hours worked generates an income effect, with additional resources coming from a boost in labor income. At the same time, changes in maternal hours worked can also generate a substitution effect, with changes in the

[6]Rapoport and Docquier (2006) and Yang (2011) both provide empirical evidence that migration increases remittances. Additionally, Akee et al. (2018) find that household financial wellbeing has large beneficial effects on children's emotional and behavioral skills.

[7]A mother may have to conduct more household production and thus spend less time with children when the father is away.

[8]Parental migration is also associated with parental absence which is found to have an adverse effect on children's psychological development and emotional well-being (Ginther and Pollak, 2004; McLanahan et al., 2013). Most existing research on parental absence is in the context of developed countries where parental absence is usually caused by divorce, military separation, or non-martial childbearing. In developing countries such as China and Mexico, parental absence is usually caused by internal or international migration (Nobles, 2011).

time that mothers allocate to childcare. Parental migration induces a similar tradeoff between family income and parental time inputs, but the empirical evidence on these mechanisms is limited.

Besides financial and time inputs, parental migration can affect child development through other important channels that are less discussed in the literature. On the one hand, parental migration can affect children's development outcomes by affecting children's own time allocation. In the context of rural families, when the father is away, a child may have to spend more time in housework or farm work instead of interacting with friends. Empirical evidence in child development literature suggests that how children spend time on their own becomes important as children grow into adolescence (see Del Boca et al. (2017); Fiorini and Keane (2014); Kooreman (2007) for examples). In the meantime, the effect of parental time investments on child development declines during adolescence (see Del Boca et al. (2017, 2014); Cunha and Heckman (2008) for examples). This decline occurs because adolescents begin to take responsibility for their own actions[9] and their developmental outcomes begin to depend on their own decisions. Although the role of children's own time investment has emerged in recent child development literature, this channel is seldom studied in the migration literature.[10]

On the other hand, many left-behind children have only one parent who is mi-

[9]Lundberg et al. (2009) distinguishes children's decisions taken on their own and those shared with their parents, and find that the probability of children making independent decisions increases sharply between the ages of 10 and 14.

[10]Antman (2011) is an exception that studies how the father's U.S. migration affect children's study and work hours.

grating, usually, the father. The father's migration can increase his income, which might increase his bargaining powers over the mother. Because mothers are more likely to spend family resources on child development including education, nutrition, and health-related commodities (Hoddinott and Haddad, 1995; Lépine and Strobl, 2013; Li and Wu, 2011), the father's migration can affect not only the amount of resources available to their children but also the amount of resources that are actually allocated to their children. Differentiating between the two is important because only the latter makes the real impact. While previous studies have emphasized the importance of intra-household bargaining in determining child outcomes, empirical studies are limited on how migration affects household bargaining and whether it is an important channel through which parental migration affects children outcomes.

In addition, we also explore parenting style (Doepke and Zilibotti, 2017; Dooley and Stewart, 2007; Fiorini and Keane, 2014) and parental mental health (Antman, 2010; Barrett and Mosca, 2013) as additional mechanisms. Both of these mechanisms have been regarded as important for child development. On the one hand, father's absence may lead to harsher (maternal) parenting style due to the lack of paternal involvement. On the other hand, father's absence could also has negative effect on maternal mental health.

To assess different mechanisms behind the effect of parental migration, we consider four different sets of outcome variables.[11] First, we consider household income and

[11]This paper does not tend to estimate a production function of children's skill formation (such as Cunha and Heckman (2008) mainly because we only have two waves of data and two measures

child-related spending. Thus, we investigate not only the available financial resources, but also how they are allocated to children. Second, we investigate the mechanism of parental time input by studying the effect of the father's migration on how many hours the father and mother spend on helping children with homework and on playing and talking with children every week. The latter could be particularly salient in the context of non-cognitive development. Third, we estimate how parental migration affects children's own time allocation among different activities. Fourth, we consider variables measuring father's decision power regarding a wide range of family decisions. These variables help us understand whether changes in intra-household bargaining is an important mechanism through which parental migration affects the allocation of financial resources to children. Finally, we also consider a proxy variable for parenting style as well as parental health as additional channels. The results shed some light on the relative importance of different mechanisms that drive the relationship between parental migration and children development.

Another important contribution of our paper is that we are able to distinguish between intra- and inter-provincial migration by looking at the information on parental migration destination. This distinction is important because it may have different impacts on left-behind children. On the one hand, the expected wage gain from inter-provincial migration is larger than intra-provincial migration. On the other hand, the loss of parental supervision and interaction might be much more detrimental for

for non-cognitive skills.

inter-provincial migration with long-distance travel and infrequent returns, than for short-distance intra-provincial migrations.[12] As we show later, intra-provincial migration generally has no significant effect on children's outcomes, while the effect of inter-provincial migration is usually significant. This is probably because the fact that intra-provincial migrants could easily travel back and forth between the workplace and home, so they are not fully absent from their children as parents who have migrated to different provinces. Existing studies on migration in China and in other countries usually do not distinguish the two types of migration.[13] By exploring this new dimension of migration, our analysis improves the understanding of the impact of parental migration on left behind children.

The main difficulty in estimating the causal effect of parental migration on children left behind is the endogeneity problem associated with migration. To be more specific, migration is not randomly assigned, and one has to worry about unobserved characteristics that are related to migration can also influence the outcome of left-behind children. For example, an unobserved negative financial shock may have a negative impact on children's development and influence household migration decision. On the one hand, since the shock makes the sending household poorer, it may encourage the father to migrate. On the other hand, if migration is costly, such

[12]With one of the few papers on this subject, Su et al. (2018) noticed that migrants who have come from other provinces differ significantly from migrants who come from the same province as the destination city in a variety of individual characteristics. In general, the migrants from other provinces tend to be younger, more likely to be male, usually in better physical condition, but have fewer years of formal education.

[13]Antman (2012) is one exception, showing the different impacts of parental domestic and international migration on children left behind in Mexico.

financial shock may prevent the father from migrating because it is unaffordable. Thus, the direction of such bias is theoretically uncertain. In addition, endogeneity may result from reverse causation, in which children's outcome is actually causing the father's migration rather than the other way around. For example, economic models of optimizing behavior suggest that the father's migration decision is likely to be responsive to their perception of children's behaviors. Thus, fathers are more likely to stay at home if their children face behavioral challenges. In these scenarios, a simple ordinary least squares (OLS) estimation may underestimate the impact that a migrant father has on children's non-cognitive development.

The main approach taken by existing studies to address the endogeneity problem has been the use of instrumental variables (IV) for migration. In particular, a large number of studies have used variation in historical migration rates at the local level to identify the effect of migration (Hanson and Woodruff, 2003; Mansuri, 2006; McKenzie and Rapoport, 2011). One concern with historical migration rate as instrument is that this variable is likely to reflect the level of economic development at the local area that would affect children of migrants directly. More recent studies have adopted economic conditions in destination areas as instruments (Amuedo-Dorantes and Pozo, 2010; Antman, 2011; Cortes, 2015; Yang, 2008). Such variables are arguably more valid since they are less likely to be correlated with local area unobservables. Section 3.2 provides a more detailed summary of the estimation methods that are adopted in the literature.

In this study, we address the endogeneity problem of parental migration by exploring the longitudinal data structure and by using instrumental variables. First, the longitudinal data structure makes it possible to use a child-level individual fixed effects model (FE) to address the possibility that some time-invariant characteristics may affect both the probability of parental migration and children's development. Second, following the literature cited above, we use labor market conditions in the destination provinces as instrumental variables.[14] Though many papers have studied the impact of migration on children left behind in the context of rural-urban migration in China, few papers have used economic condition in destination areas as instruments. This probably happens because the commonly used datasets lack information on migration destination.

We construct two scores to measure children's non-cognitive skills based on a serious of questions provided in the GSCF regarding their internalizing and externalizing behavioral challenges. The two scores are standardized to have a mean of 0 and a standard deviation of 1. Our estimates show significant adverse effects of the father's inter-provincial migration [15] on the non-cognitive skills of left-behind children, which reduces the scores by 0.03 and 0.06 points in regards to internalizing and externalizing behaviors, respectively. However, we find that the effects of intra-provincial

[14]Rural migrants tend to work in the manufacturing and construction sector so that changes in wage and employment levels in those sectors are more likely to affect migration behavior but are unlikely to affect children's development directly. As a robustness check, we use wage and employment in these sectors as instruments.

[15]We focus on fathers' inter-provincial migration because in our sample fathers are much more likely to migrate than mothers and intra-provincial migration has not been found to significantly affect child outcomes.

migration, though still negative, are much smaller and insignificant. Effects are also stronger for boys than girls. These results are robust with different specifications for our regression model, alternative instruments, and an alternative value-added model. Our results suggest that the inter-provincial migrations, which are quite common in China, is a much more serious problem in shaping the emotional well-being of children left behind, and therefore deserves greater policy attention. Further discussion on mechanisms shows that fathers' migration is associated with an increase in family income, but not child-related spending probably due to the increasing bargaining power among fathers. More important, fathers' migration reduces the time parents spend talking and playing with their children. These two mechanisms might be the major driving forces behind the negative effect of the father's migration on left-behind children.

The remainder of the study is organized as follows. The next section provides background information. Section 3.3 presents the empirical method. Section 3.4 introduces the data and provides summary statistics. Section 3.5 reports estimation results and robustness checks. Section 3.6 provides a comprehensive discussion of the potential mechanisms. Section 3.7 discusses sub-sample results. We conclude with Section 3.8.

3.2 Background

3.2.1 Parental Migration and Development of Children Left Behind

There are almost 1 billion migrants worldwide, with 214 million international migrants and another 740 million internal migrants moving within countries (United Nations, 2013). The impact of migration on the school attainment of children left behind is one of the most intensively studied dimensions of the consequences of parental migration. The literature so far yields ambiguous predictions. On the one hand, a number of papers find that migration increases education attendance or enrollment (Acosta, 2011; Antman, 2012; Binci and Giannelli, 2018; Cox-Edwards and Ureta, 2003; Gibson and McKenzie, 2014; Hanson and Woodruff, 2003; Mansuri, 2006; Yang, 2008). On the other hand, a substantial number of studies find that migration negatively impacts school attendance or enrollment (Amuedo-Dorantes and Pozo, 2010; Cortes, 2015; Giannelli and Mangiavacchi, 2010; Kroeger and Anderson, 2014; McKenzie and Rapoport, 2011). Empirical evidence on the effect of migration on children's cognitive skills measured by test scores is relatively limited and is far from conclusive. For example, Macours and Vakis (2010) find that parental migration increases children's academic performance in Nicaragua, while Powers (2011) find parental migration decreases children's test scores in Mexico.

More important, to the best of our knowledge, no paper so far has employed a rigorous estimation method to study the impact of parental migration on non-cognitive skills of children left behind, although a small number of papers have found correlations between parental migration and psychological well-being of children left behind. For example, in Romania, Botezat and Pfeiffer (2014) find that children's psychological well-being measured by whether they were bullied or were involved in conflicts is negatively associated with parental migration.[16]

The endogeneity of migration is one of the most important reasons that existing studies report mixed results regarding the impact of migration on children left behind. Researchers have applied various methods to address this problem. The main approach has been the use of instrumental variables for migration. The most commonly used instrumental variables can be categorized into two groups. The first group consists of a large number of studies that have used variation in historical migration rates at a local level to identify the effect of migration.[17] The second group of more recent studies has used economic conditions in destination countries or regions as instruments, to better satisfy the validation condition. For example, Yang (2008) uses variation in exchange rate appreciation in migrants' destinations to show that remittances result in an increase in child schooling. Amuedo-Dorantes and Pozo (2010) make use of unemployment and wages in U.S. destination states to identify the effects

[16]Instead of looking at left-behind children, a few papers have studied the impact of migration on mental health or psychological well-being of elderly parents left behind (Antman, 2010) and of migrants themselves (Atella et al., 2019; Barrett and Mosca, 2013; Stillman et al., 2015).

[17]For examples, see Hanson and Woodruff (2003); Mansuri (2006); McKenzie and Rapoport (2011).

of remittances on expenditure patterns in Mexico. Antman (2011) uses employment condition in U.S. cities the potential migrant could select as a destination as instruments and studied the effect of the father's migration status on children's school and work outcomes in Mexico. Gibson and McKenzie (2014) use the demand shocks in destination countries to study the effect of the mother's migration on children left behind in the Philippines.

Recently, a small number of papers have used randomized trials to study the impact of migration on the household members who are left behind. Gibson et al. (2011a) study a migration lottery program in New Zealand that allows them to compare families of successful applicants who were authorized to migrate with families of applicants that were not allowed to migrate. Mergo (2016) uses the Diversity Visa lottery program to evaluate the effects of international migration on families left behind in Ethiopia. Although the experimental approach leads the way to promising solutions for the endogeneity problem, it is usually very costly and, thus, not easily applied to alternative contexts.[18]

Besides the endogeneity problem of migration, another major challenge of the existing literature is to identify the different mechanisms that drive the effect of parental migration. So far, the empirical evidence on the mechanisms of how parental migration affects child development is limited mainly because most socioeconomic surveys

[18]A few studies have applied the matching methods, which is not usually preferred because the method has to assume that selection into migration depends on observable characteristics only and match migrants with comparable non-migrants based on those observable characteristics. For example, see Esquivel and Huerta-Pineda (2007).

lack precise measures of parental investments and other factors that are important in the production of children's skills. Nevertheless, the existing literature has provided some discussions on how parental migration affects material inputs and children's own time allocation.

First, many studies have used household income or remittance as a proxy measure for material inputs in understanding the impact of parental migration on child development. A consensus in this strand of literature is that migration increases remittances and household income (Alcaraz et al., 2012; Cox-Edwards and Ureta, 2003; Gibson and McKenzie, 2014; Yang, 2008). The remittances can ease household budget constraint and, thereby, can potentially increase household spending on children. Thus, parental migration can have a positive impact on children's development. However, increased household income may not always lead to an increased spending on children. For example, if the father's migration increase his bargaining power and he chooses to spend less on children, then the father's migration can reduce financial resources allocated to the children. The richness of our data allows us to not only look at the overall income but also child-specific expenditure on different categories.

Moreover, in the literature of child development, parental time inputs are considered as one of the most important determinants of child outcomes, however, we know relatively little about how migration affects the time allocation of parents. Though it is reasonable to assume that a parent who migrates will spend less time with his or her children, it is possible that the other parent who is left behind with the chil-

dren will sacrifice leisure time to compensate for the reduced time inputs. The effect of migration on parental time inputs is seldom studied in the literature. Zhang et al. (2014) is one exception, providing summary statistics showing that families with migrant parents in China are less likely to tutor their children. [19] Our paper goes beyond the current literature by providing causal evidence of parental migration on parental time inputs.

In the meanwhile, a few papers have studied how parental migration affects children's own time allocation. For example, McKenzie and Rapoport (2011) provides summary statistics showing that children with migrant parents are more likely to devote time in housework and family business. Antman (2011) studies how parental U.S. migration affects the number of hours per week that a child devotes to studying and working and finds that parental migration reduces study hours and increases work hours. Chen (2013) examines the time allocation of left-behind children in China and finds that children spend more time in household work. Migration costs can also directly affect children's school decision as shown in de Brauw and Giles (2017). They find a negative relationship between migrant opportunity and high school enrollment. With rich information on household spending, intra-household bargaining, and parent and child time allocation, our paper contributes to the literature by looking at the mechanisms behind the link of parental migration and child non-cognitive skills in a more comprehensive and coherent way.

[19] Zhang et al. (2014) also shows that families with migrant parents less likely to have a satisfying relationship between children and parents or other adults in the family.

3.2.2 Labor Migration and Children Left Behind in Rural China

Beginning in 1958, China started the household registration system (*hukou*) which categorized citizens into rural or urban residents and imposed strict restrictions on rural residents migrating to urban areas.[20] These restrictions were only gradually relaxed by the government in the late 1980s when China's open-door economic reform led to the rising demand of labor in urban areas, while, concurrently, privatization of farming let to a surplus of labor in rural areas. Since then, hundreds of millions of rural migrant workers have travelled to cities to find jobs. According to a survey from the National Bureau of Statistics of China, the total number of rural migrant workers has grown from 104 million in 2002 to 168 million in 2015. Such a huge wave of rural-to-urban migration is unprecedented and has been described as the largest peacetime migration in human history (Du et al., 2005; Zhao, 1999).

Rural migrants are allowed to work and live in urban areas as temporary residents. However, they are not able to access the urban welfare system, including education health, and the social safety net, mainly because of the current *hukou* system. The majority of these migrant workers perform low-paying jobs and live in crowded conditions (World Bank, 2009). As a result, a considerable number of migrant parents (often migrant father) choose to leave their children and spouses behind in their rural

[20]See Song (2014) for a detailed introduction of the origins and changes of the *hukou* system in China.

communities.[21] Recent estimates based on the 2010 population census suggest that there are more than 61 million children aged 17 years or below that have been left behind by their parents in rural areas, which account for 37.7 percent of rural children and 21.9 percent of all children in China (Zhang et al., 2014).

This left-behind children phenomenon has attracted increasing attention from both policymakers and academic researchers. So far, despite the empirical evidence from other countries that produces very mixed results, a majority of studies in China find negative impacts of parental migration on educational outcomes of children left behind. To be more specific, these studies find that parental migration reduces school attendance (Hu, 2012; Meyerhoefer and Chen, 2011; Wang, 2014) and decreases test scores (Meng and Yamauchi, 2017; Zhang et al., 2014; Zhao et al., 2014). In contrast, Bai et al. (2017) finds that parental migration has a positive impact on children's academic performance.

More relevantly, some work has studied the effect of parental migration on children's emotional or psychological well-being, and the evidence on this topic is far less conclusive and remains mainly descriptive. Some studies suggest that left behind children are more likely to have symptoms of anxiety and depression (Liu et al., 2009) and are more likely to feel sad and think of suicide or leave home (Gao et al., 2010). Other studies find that these impacts are insignificant (Ren and Treiman, 2016; Xu and Xie, 2015).

[21]Despite an increasing number of rural children accompanying their parents and entering the migration process, single person migration remains the dominant pattern in the internal migratory flow in China (World Bank, 2009).

Thus, this paper makes two major contributions to the literature. First, it expands the scope of the current literature by studying the impact of parental migration on left-behind children in terms of non-cognitive skills, which is seldom studied in the literature in either China or other developing countries. More important, this paper improves the identification of the causal effect of parental migration in China by adopting economic conditions in destination areas as instrumental variables. As described section 3.2, these variables are commonly employed in the literature and have been proven to be valid instruments (**?**). However, to the best of our knowledge, not many papers on left-behind children in China have explored these variables.

The most similar paper to ours is (Lee and Park, 2010) which also uses the GSCF to study the effect of parental migration on a variety of measures of child development, including school enrollment, years held back, achievement scores, and psychological well-being. In particular, they find no impact of parental migration on children's internalizing behavior while some suggestive evidence on the negative impact on children's externalizing behavior. Our work differs in several important ways. First and foremost, we emphasize the underlying mechanisms through which parental migration affects child development, which are not analyzed in Lee and Park (2010). Understanding the mechanisms are particularly important in the context of parental migration, because the directions of various channels mentioned above are different. Having a better understanding of the underlying mechanisms is particularly important for policy recommendation. Second, we define a father's migration status by his

workplace location which allows us to differentiate between inter- and intra-provincial migration. This differentiation is important because different types of migration lead to different impacts. Lee and Park (2010) defines a father as migrant if he is away from home for more than 3 months in the past year. Although this definition captures the fact that a migrant father may be away from home for a longer time, it is not clear whether the father is away consecutively or if the father travels back for the weekend.[22] Thus, that definition of migrant might be ambiguous if we can not differentiate between consecutively away and only away during weekdays. Moreover, our identification strategy is improved and has stronger identification power.[23] Finally, we also refine the measurement of children's non-cognitive skills using item response theory. Thus, our results could complement findings in Lee and Park (2010) by improving our understanding of how parental migration affects child development.

3.3 Empirical Strategy

Since our primary goal is to estimate the effect of the father's current migration on his children's non-cognitive development, the simplest econometric framework might

[22] In both cases, fathers are defined as migrants, but the impact on their children might be quite different. In the latter case, fathers would still be able to communicate with their children on a regular base, while in the former case, communication is less likely to happen.

[23] Lee and Park (2010) has weak instrumental variable restrictions.

begin by estimating the following equation:

$$Y_{it} = \beta_0 + \beta_1 fm_{it} + \beta_2 X_{it} + \epsilon_{it},^{24} \tag{3.1}$$

where Y is the non-cognitive scores for child i at time t, which are measured for internalizing as well as externalizing behaviors. The effect of interest is captured by the coefficient on the fm variable, which is an indicator that equals 1 if the father migrates out of Gansu province and 0 otherwise. This means that the reference group in the analysis includes children with non-migrant fathers and children whose father migrates within the province. Moreover, the binary definition of migration status also makes it possible to adopt destination shocks as instrument variables, which could help us better deal with the endogeneity problem of migration, while including intra-provincial migration may require additional instruments that are not available.[25]

The vector of covariates X_{it}, includes a rich set of control variables. At the child level, we control for a child's age and gender. At the parental level, we control for a father's and mother's age and education level. We include both parents' education levels instead of just that of the household head to better capture children's inheritable ability. More important, previous research has shown that a mother's education has a stronger impact than a father's education on a child's developmental outcomes (Dickson et al., 2016). At the household level, we control for the number of siblings

[24]The linear regression is applied because the non-linear models cannot accommodate the combined instrument variable and fixed effects estimation (Antman, 2011),

[25]Many papers on children's cognitive and non-cognitive development have used a production approach using a value-added model (Todd and Wolpin, 2003, 2007). Nevertheless, we adopt a value-added model as a robustness check. Our main results do not vary much with different model specification.

and whether the family lives together with at least one of the grandparents. The variables are important factors determining the available time and financial resources and how these resources allocated within a household. In addition, we include county fixed effects to avoid bias from omitted local-level factors.

As discussed, one concern with estimating equation (1) is that the OLS estimation method will yield biased estimates of β_1 since the fm_{it} variable is endogenous. First, fathers and children may share some personality traits that are persistent over time, and those traits may affect a father's migration decisions and children's behaviors simultaneously. The panel nature of the data allows a simple solution to correct for this type of endogeneity by adding child fixed effects. Thus, the regression model can be expressed as following:

$$Y_{it} = \beta_0 + \beta_1 fm_{it} + \beta_2 X_{it} + \eta_i + \epsilon_{it}, \tag{3.2}$$

where η_i captures time-invariant observed and unobserved heterogeneity at the individual child level. However, there could still be some sources of endogeneity that vary over time. For example, a negative shock to the family's income may affect children's psychosocial development, while at the same time, this shock could force the father to migrate for higher income or prevent the father from migrating by making migration unaffordable. In this case, the direction of such bias is theoretically uncertain. Moreover, parents may respond to realized poor behaviors by staying at home to accompany their children. The OLS estimate will be biased upward if we fail to control this bias.

To address this concern, we propose a set of instrumental variables that will only influence the child's outcomes through their effects on the father's migration status. The proposed instruments are based on labor market conditions in the destination provinces where the father was most likely to migrate to.[26] Economic conditions in the recent past in the destination provinces can affect the father's migration decision without influencing the children left behind directly.

The main empirical strategy then amounts to the estimation of equation (2) above by instrumental variables, where migration status is estimated through the following first stage regression:

$$fm_{it} = \alpha_0 + \alpha_1 Z_{it} + \alpha_2 X_{it} + \mu_i + \epsilon_{it}, \tag{3.3}$$

where Z_{it} is a vector of instrumental variables excluded from equation (2). The set of variables Z_{it} comprise of the employment and wage in the destination provinces to which the father was likely to migrate. Since the child fixed effects model with instrumental variable (FEIV) approach involves using repeated observations of children from the same family in different time periods, we cluster the standard errors at the level of the individual to allow for arbitrary correlation within the individual and across time.

[26] A growing number of papers have used economic conditions in the destination areas as instruments to estimate the effect of migration. For example, see Amuedo-Dorantes and Pozo (2010); Antman (2011); Yang (2008),and Cortes (2015).

3.4 Data and Summary Statistics

This study analyzes data from the GSCF, a longitudinal study of 2,000 children who were 9-12 years old in the year 2000 and living in 100 rural villages in Gansu. The data is designed so that such 2,000 children constitute a random sample of all children in this region. Gansu is located in the northwest of China (see Figure 3.1) and is one of the poorest provinces in the country. According to the National Bureau of Statistics (NBS), in 2001 and 2005, rural per capita income in Gansu ranked 30th out of 31 provinces.

This study uses data from the first two waves of the GSCF, which were conducted in 2000 and 2004.[27] The GSCF has a low sample attrition because 1,872 (93.6 percent) children of the original 2,000 children were re-interviewed in wave 2 (2004). In each wave, the GSCF collected detailed information for the target children on their developmental outcomes including cognitive and non-cognitive skills, as well as parental, household, and environmental factors that may affect these developmental outcomes. In particular, it also contains information that helps define father's migration (discussed below) as well as various mechanisms mentioned above. Specifically, the data contains different sources of income including remittances, information about parental and children's own time allocation, as well as measures of household

[27]Wave 3 of the GSCF was conducted in 2007-2009 when sample children were 17-21 years old. Since the impact of parental input becomes less important in shaping their non-cognitive skills (Cunha et al., 2010) at those ages, we choose to use only waves 1 and 2. Moreover, wave 3 covers a highly selective sample because its data collection process is different from previous waves. See Glewwe et al. (2017) for a more detailed explanation on this issue.

bargaining power.[28]

We impose a set of sample restrictions. First, given that the mother's migration is rare in our sample,[29] we focus on the impact of the migrant father and restrict our sample to children whose mother stayed at home. Second, we exclude children whose parents are divorced or deceased to eliminate the confounding factors associated with the family structure.[30] In addition, to use a child fixed effects model, we exclude children that appear only in wave 1. We further drop observations with missing values of main variables. Our final data consists of 1,779 children with 3,558 observations.

The GSCF provides a unique data source that enables us to study the development of children's non-cognitive skills in the context of a developing country.[31] It asks the respondent child a series of questions regarding their *internalizing* and *externalizing* behavioral challenges. The former captures intra-personal problems such as withdrawal, depression, and anxiety, and the latter captures inter-personal problems such as destructive behavior, aggression, and hyperactivity. We construct two variables measuring children's non-cognitive skills that capture internalizing and externalizing behaviors separately. Each measurement is constructed by applying item response

[28]Because of its richness on child development, the GSCF has been widely used to analyze how parental and school inputs affect children's educational achievements, health outcomes, and non-cognitive behaviors. See Glewwe et al. (2016, 2017); Hannum and Zhang (2012); Leight (2017); Leight and Liu (2016) for examples.

[29]In the GSCF, less than 3 percent of children have a migrant mother. Using the same data, Lee and Park (2010) reports the same migration rate for mothers.

[30]Compared with the United States, divorce is relatively uncommon in China. In 2000, the U.S. divorce rate (number of divorces per 1,000 population) was 4 according to the census, while China's divorce rate was only around 1 according to the China Statistical Yearbook.

[31]See Almlund et al. (2011); Borghans et al. (2008); Humphries and Kosse (2017); Thiel and Thomsen (2013) for measurements and interpretations of non-cognitive skills across different disciplines.

theory (IRT) to generate a single index based on a list of 17 questions (Andrich, 1978). We use the IRT approach to measure children's non-cognitive skills.[32] An example of a question used for the internalizing index is whether you fully agreed, agreed, disagreed, or totally disagreed with the following statement: "I easily get anxious." An example of a question used for the externalizing index is whether you fully agreed, agreed, disagreed, or totally disagreed with the following statement: "I often lose my temper with others." A full list of questions is summarized in Appendix 3.14. The two scores are standardized to have a mean of 0 and a standard deviation of 1. A higher score means fewer behavioral problems and higher non-cognitive skills.[33]

Throughout this paper, we define a father's migration status by the location of his workplace. In particular, we define a father to be a non-migrant if his workplace is located in the village, an intra-provincial migrant if his workplace is located outside of the village but within Gansu province and an inter-provincial migrant if his workplace is located in another province outside of Gansu.[34] As we plan to adopt the child

[32]IRT is a well-established method that the literature use to aggregate information from different measurements. It has several advantages than simply summarizing answers to different questions. Each question provides different information about the underlying non-cognitive skills that are being measured. There is no reason to believe that a score of 4 in one of the measure implies the same level of non-cognitive development as a score of 4 in another measure. For example, fully agreeing that "I steal things from others" might indicate something different than fully agreeing that "I often make fun of others." Summing up the scores ignores this issue. If we think that each question is measured with some noise and that the variance in the noise is different across questions, then summing up the scores on each question will provide a very unreliable and noisy measure of the underlying non-cognitive skills.

[33]Glewwe et al. (2017) and Leight and Liu (2016) constructed measures of non-cognitive skills the same way.

[34]Previous papers usually define a migrant parent if he or she is away from home for certain period of time during the year. See Lee and Park (2010) and Zhang et al. (2014) for more information. As illustrated in Table A2, our definition is consistent with that in the literature as parents with a workplace outside of the village experienced much longer periods of time away from home.

fixed effects model, it is important to have enough over-time variation in the father's migration. We observe this variation in panel A of Table 3.1, which categorizes our sample children by their father's migration status in 2000 and 2004. In 2000, 62.9 percent of fathers stayed in the home village, 30.8 percent had previously migrated within the province, and 6.3 percent had previously migrated out of the province. In 2004, 65.1 percent of fathers stayed in the home village, 26.5 percent had previously migrated within the province, and 8.4 percent had previously migrated out of the province. This trend is consistent with the growing migration rate at the national level (Su et al., 2018).[35]

A further look at the longitudinal structure of our data shows a large variation of the father's migration over time within a family (panel B of Table 3.1). The majority of fathers never migrated out of the province (88.5 percent). There are 2.6 percent of fathers migrated out of the province in both 2000 and 2004, 3.7 percent migrated out of the province only in 2000, and 5.2 percent migrated out of the province only in 2004.

Table 3.2 shows the summary statistics for children in our final sample, as well as three subsamples defined by the father's migration status.[36] A clear relationship

[35]In Table 3.15, we report the amount of months the father has been away from home in each category by years, showing that inter-provincial migrants have been away from home for a much longer time than intra-provincial immigrants (6 months versus 3 months) in 2000. In wave 2, the GSCF asked two different questions on how many months the father has been away. The first question asked how many months the father has been away consecutively, and the second question asked how many months the father has been away only during weekdays. As shown in Table 3.15, inter-provincial migrants are away for more months both consecutively and during weekdays than intra-provincial migrants (4.5 and 0.9 months versus 2.0 and 0.7 months).

[36]In Table 3.16 and Table 3.17, we present descriptive analysis for children in 2000 and 2004 separately.

emerges between the father's migration status and non-cognitive skills. Children with non-migrant fathers perform better in both internalizing and externalizing behaviors (higher scores for better non-cognitive skills) compared with children with migrant fathers. Among children with migrant fathers, those with fathers who migrated out of the province have the lowest non-cognitive scores. To be more specific, average non-cognitive scores (with mean 0 and s.d. 1) of children with non-migrant fathers range from 0.019 to 0.003. Such scores on average range from -0.014 to 0.013 for children with an intra-provincial migrant father, and on average range from -0.115 to -0.078 for children with an interprovincial migrant father. That is, the farther away the father was from home, the lower scores the children had.

In terms of control variables, a first look at the statistics for all children (column (1) of Table 3.2) reveals that the average age of sample children is 13 years old and 47 percent of them are female. On average, fathers and mothers in our sample are 39 and 37 years old, respectively. Fathers had 6.7 years of education, while mothers had about 3.7 years, which suggests that most fathers do not finish junior middle school and most mothers barely finish elementary school. Families in our sample on average have 2.3 children, which suggests that the one-child policy was not strictly enforced in the rural area of Gansu. About 20 percent of families have at least one grandparent living with them. Regarding village level characteristics, the average village population in our sample is 1712, and the average arable land per capita is 2.2 mu (about 8.1 square meters). About 32.9 percent of villages are in plain terrain, 9.8

percent in hills terrain, 33.9 percent in mountain terrain, and 23.4 percent in other terrains. A majority of villages have their own elementary schools (83 percent), while a relatively smaller number of villages have their own middle schools (10 percent).

Columns (2)-(4) of Table 3.2 report the summary statistics across the three sub-samples according to the father's migration status, showing that most demographic characteristics across the three groups do not vary much. However, we do find some interesting differences, especially between the samples with an inter-provincial migrant father and other groups. For instance, fathers and mothers in the inter-provincial group have, on average, 5.9 and 2.4 years of education, which are the lowest among the three sub-samples. They also have more children and are more likely to have at least one grandparent at home. These characteristics are consistent with what has been found in Su et al. (2018).

Our instrumental variables are economic conditions in the recent past of destination provinces where the father is most likely to migrate. We identify the provinces to which fathers in our sample are most likely to migrate based on the village module of GSCF. The village module of GSCF asks the top migration destination province of the village, and is answered by the village leader. One of the main reasons for us to use this village level information is due to the fact we only know whether a father migrates or not, we do not know which geographic area, e.g. which province, the father has moved to at the individual household level. Hence, our instrument is at the vil-

lage level.[37] As reported in Table 3.3, the top migration destinations for each village appeared to be relatively stable and are concentrated in certain provinces during our sample period. The most popular destination is Xinjiang province, which accounted for about 37 percent of the top migration destination in 2000 and 39 percent in 2004. Xinjiang is to the northwest of Gansu, but unlike Gansu, Xinjiang's economy has been performing close to the national average ever since the mid-1990s, owing perhaps to the western development policies launched by China's central government. One of the important goals of these policies is to enhance national unity and social stability by facilitating economic and social development in the west Goodman (2004).[38] In 2000, the average annual wage was 7,605 RMB in Xinjiang and 7,277 RMB in Gansu. The economic differences between Gansu and Xinjiang explain why there is a mass migration from one province to the other. This migration trend is consistent with what has been observed in studies on inter-migration in China (Su et al., 2018). Other common destinations include provinces near Gansu (Ningxia, Qinghai, and Shaanxi) or provinces in booming coastal regions (Guangdong and Beijing).[39]

Once we have identified the destination provinces to which migrant fathers are most likely to move, we link the child observations with employment and wage data

[37]Such migration destination provinces are likely to be formed as a result of historical and geographical factors.

[38]Xinjiang has received particular attention from the central authorities because it borders Central Asia and has experienced a growth of separatist movements. To contain separatism, China's central government has stepped up its efforts to improve living standards in Xinjiang and to crack down on separatism. Fiscal transfers from the central government to Xinjiang have increased from 5.91 billion RMB in 1996 to 18.4 billion RMB in 2001 (Information Office of the State Council 2003).

[39]The GSCF also asked the migrant father which province his workplace was located in if his workplace was outside of Gansu. However, only about half of our observations have valid answers to this question.

in the most likely destination province. The province-level employment and wage data are available from the China Labor Statistical Yearbook for 1999 and 2003. It is expected that these variables will act to stimulate migration. That is, potential migrants will be more likely to make the trip when employment and wage are high indicating a boom in those industries. As a robustness check, we also construct the instrumental variable using employment and wages in the manufacturing and construction sectors, separately.[40] Finally, given the child fixed effect, change of values in our instrumental variable across two waves is also needed to achieve identification.

To investigate the potential mechanisms of how the father's migration affects children's outcomes, we explore variables that measure household financial inputs, household bargaining power, and parental and children's time inputs. In particular, we measure household financial inputs by household income, household expenditure, and household item purchase. Table 3.4 panel 1 reports household income across different groups of households and shows that the average total household income (including remittances) is 10,121 RMB, which consists of about 62 percent (6,311 RMB) agriculture income and about 27 percent (2,721 RMB) labor income. Overall, the three groups of households have a similar level of household total income, but with quite different components. As expected, households with a migrant father have a much higher labor income (5,815 RMB, 4,854 RMB, and 1,412 RMB), but a much lower agriculture income (3781 RMB, 5387 RMB, and 7015 RMB) than households

[40]Manufacturing and construction sectors employ a large number of migrant workers, but the employment and wage variables in those sectors are highly correlated with those at the aggregate level, so we construct our instrument based on aggregate variables.

in which the father stays at home. Table 3.4 panel 2 reports, the annual expenditure on children's clothes, school supplies, and children's allowances are 234 RMB, 35 RMB, and 93 RMB, respectively. When we look at these expenditures by migration status of fathers, a clear trend emerges: households with a migrant father, especially an inter-provincial migrant father, have much lower spending on children. Table 3.4 panel 3 describes a similar story as in panel 2, showing that households with an inter-provincial migrant father are less likely to purchase reading materials, a dictionary, and desks for their children. Considering that different groups of households receive similar total household income (Table 3.2), financial constraints do not seem to be the main explanation of why households with migrant fathers spend less on their children.

Table 3.5 reports parental time inputs which are measured by weekly hours the father spends with his children. For a father who is absent from home, those variables reflect weekly hours he spends on those activities when he stays at home. We consider two variables: weekly hours helping children with their homework and weekly hours of playing or talking with children. Table 3.5 shows that even when they come back home, fathers with inter-provincial migration experience spend much less time playing and talking with their children (2.5 hours) compared with fathers who stay home (2.9 hours) and fathers who migrate within Gansu (2.8 hours). This might occur because when migrant fathers come home from other provinces, they have to take in charge of certain home production activities, such as house repairs, that are usually conducted

by male members of the households. We indeed observe that inter-provincial migrant fathers spend about 12 hours on home production, while nonmigrant fathers and fathers migrant within Gansu only spend about 10 or 7 hours, respectively. This further implies, when fathers are away, mothers may have to undertake some of the home production that the faters were doing, which is also confirmed below.[41]

A father's migration can also affect a mother's time inputs. Table 3.5 shows that on average mothers in our sample spend 30 hours on home production, 1.3 hours helping their children with homework and 3.7 hours playing and talking with their children every week. If the father migrated out of Gansu province, the mother will spend 4 more hours on home production comparing with the mother whose spouse stays at home. In the meantime, mothers will spend fewer hours with their children, especially helping their children with homework. Overall, summary statistics in Table 3.5 suggest that a father's migration status largely reduces the number of hours a father and mother can spend with their children, which is likely to be the major mechanism through which a father's migrant status affects children's development.

In the bottom panel of Table 3.5, we report children's time allocation measured by their weekly hours spent on home production, taking care of others, doing homework, and watching TV. Overall, we find that children with a migrant father are more likely to participate in home production and caring for others, while the difference in doing homework and watching TV is not evident.

[41] Although we can not observe how many hours migrant fathers spend with their children when they are away from their residents, we reasonably assume that the time they spend with their children is quite limited because of their absence.

Table 3.6 reports the father's decision power relative to the mother, which is measured by the questions regarding whether the father is the decision-maker behind a wide range of family decisions. These decisions include the children's schooling, how to treat the child, purchase of durable goods, management of family finances, planting of crops, and livestock transaction. Those questions are answered by the mother. If the father makes those decisions, the variables are coded as 1; if the mother makes the decision or if the parents make the decision together, the variables are coded as 0. Thus, the variables reflect the bargaining power of the father. Overall, the father is not the major decision maker in the household. On the management of the family finances, fathers have about a 42 percent probability of being the decision maker, which is the highest among all decisions asked in the survey. Meanwhile, the father's migration status does not seem to cause large differences according to the summary statistics presented here.

3.5 Father's Migration and Children's Non-cognitive Skills

3.5.1 Main Results

Table 3.7 presents our main estimation results. In columns (1) and (5), we start with the simple OLS estimation, which assumes the father's migration status is exoge-

nous. The coefficient estimates on the dummy indicator for having an inter-provincial migrant father are negative and significant for both non-cognitive scores. In particular, having a father who migrates out of the province is associated with a 0.046 and 0.084 declines for internalizing and externalizing scores, respectively, compared to having a father who stays at home or migrate within the province.

The simple regression coefficients presented in columns (1) and (5) are subject to the potential endogeneity problem. As a first attempt to address this issue, we employ a child fixed effects model to deals with the time-invariant source of endogeneity. Columns (3) and (7) report the estimation results of the fixed effects model. The estimates of the coefficient on inter-provincial migration remain significant and are consistently smaller in magnitude than in the linear regression estimates reported in columns (1) and (5) (-0.025 for internalizing behaviors and -0.054 for externalizing behaviors), which suggest that the inter-provincial migration is positively correlated with the omitted time-invariant determinants of a child's non-cognitive skills. This is consistent with the notion that migration households tend to pay less attention to children's development or are more financially distressed.

A fixed effect model helps to deal with the time-invariant omitted variable problem; however, the above results may still be contaminated by time-varying confounding factors and reverse causality problem. We deal with these problems with the instrumental variable method. Table 3.7 columns (2) and (6) report the estimation results for two-stage least square estimation and columns (4) and (8) report estima-

tion results for the fixed effects model with instrumental variables. Consistent with the OLS and fixed effects model, the estimation coefficients for inter-provincial migration are significantly negative, -0.058 and -0.029 for the internalizing score and -0.088 and -0.058 for the externalizing score. More important, the estimated coefficients for father migration become larger after adopting instrumental variables in both models, which suggests that time-varying unobservables, such as negative income shocks, tend to prevent migration and bias down the estimation. Nevertheless, it is worth noting that the instrumental variable has a much smaller impact on father's migration coefficient than the child fixed effect does. The overall effects of a father's migration appear to have negative impacts on children's non-cognitive development. In terms of the mechanism behind this negative effect, we test the financial inputs, time inputs, and other related channels in later sections.

The first stage regression results are presented in Table 3.8 using both OLS and fixed effects models. Employment and wage levels in the top destination provinces are included together as instrumental variables. Both employment and wage levels are one year ahead of the survey year. As observed, both instruments are good predictors of the likelihood that the father's migration out of province. The point estimates in column (2), where we adopt the fixed effects model, indicate that an increase in employment by 1,000 would correspond to an increase in the probability of a father migrating by 0.02 percent and an increase in wage by 1,000 RMB would increase the probability of paternal migration by 1.4 percent. The coefficient is more significant

for destination wage, which is at 1 percent level. As indicated at the bottom of Table 3.7, the F statistics on the weak instrument test ranges from 17 to 35 across different specifications, which are well above the critical values for weak instruments as reported by Stock et al. (2002a). In addition, since there are two instrumental variables, an overidentification test is also possible, although it can be argued that if both instrumental variables are measuring the same economic forces, the test provides limited information (Murray, 2006; Sargan, 1988). Nevertheless, we fail to reject the null hypothesis of valid instruments according to the overidentification test.

In Table 3.9, we consider two alternative sets of instrumental variables by exploring employment and wage in the manufacturing and construction sectors which are most likely to hire migrant workers. The estimation results are quite similar to our main results, with F statistics on the weak instrument test ranging from 15 to 34 across different specifications. Table 3.8 columns (3)-(6) report the first-stage estimations using employment and wage in the manufacturing and construction sectors as instrumental variables for the father's migration. Estimates are qualitatively similar to our baseline case in columns (1) and (2).

3.5.2 Robustness Check

Here we discuss the robustness of our results considering several possible alternative model specifications. First, one concern to our identification strategy lies in the exclusion restriction necessary for the instrumental variables estimation. It is

possible that destination economic conditions affect child outcomes directly, perhaps because the national-level business cycle will affect both the destination province and the local economics, which will certainly affect the child outcomes. Another possible threat to the validity of our instruments could come from the fact that migrants from different regions may traditionally send migrants to specific provinces. In that case, the instruments could be simply capturing differences across regions, such as differences in overall economic development. To address these concerns, we include the county and the year interaction term in Table 3.18 columns (1) and (6). Our baseline results are robust to the inclusion of those variables.

In addition, the following sets of variables have been found to be relevant to explain children's non-cognitive development in previous papers: (1) children's cognitive skill and health status, (2) school inputs, and (3) parent health status. In Table 3.18 columns (2)-(4) and (7)-(9), we begin by considering the subsample of children for whom we can observe health status and cognitive skills. Health status is self-reported and range from 1 to 4, with the higher scores indicating better health. The cognitive skills are measured by math and language scores based on standardized tests developed for the survey by test experts in the Gansu Educational Bureau. The test scores are normalized by standard deviations from the mean score of children in the same grade level. We do not observe children's cognitive skills for the full sample because only half the students were given language tests and the other half were given math tests in 2000. Nevertheless, the magnitude of the migration coefficients is very similar

to that observed in our main estimation (columns (4) and (8) in Table 3.7), despite the increased standard errors caused by the smaller sample size.

In columns (3) and (8) of Table 3.18, we test for the omission of school inputs, using information about schools that are self-reported by the sample children.[42] These variables describe whether the sample children observe disruptive behaviors in the school, such as violent discipline, disruptive classes, absentee teachers, and school closing for no legitimate reasons. Again, we do not observe any change in our coefficients of interest. In columns (4) and (9), we check for a potential bias caused by the omission of parent's health shocks by including two variables that measure parent's health status. The two variables are self-reported and range from 1 to 4, with higher scores indicating better health. Estimation results show that our main results are robust to the inclusion of these measures of parents' health conditions. Overall, we conclude from our sensitivity analysis that our results are relatively invariant with respect to these changes in the model specification.

In addition, household inputs such as purchased goods and services as well as time inputs may also affect children's non-cognitive development. As mentioned by Todd and Wolpin (2007), one way to account for missing variables on such inputs is to approximate family income and parent allocation of time. However, their inclusion is problematic because they will be affected by the father's migration, making it hard to

[42]The GSFS has a separate questionnaire which covers more detailed information on schools that were answered by the school principal. However, since a large proportion of observations (about 70 percent) in the household survey are missing school ID, merging variables from this questionnaire to our sample will dramatically reduce the sample size. Thus, we chose not to use this questionnaire.

interpret our estimation results (Bono et al., 2016; Ermisch and Francesconi, 2013). Nevertheless, in addition to the variables used in the benchmark specifications, we include a log of family annual income and the father's and mother's weekly hours spent with the children as control variables in columns (5) and (10) of Table 3.18. Regardless of the outcome, the estimates from the specification with added controls are remarkably similar to our baseline results. This provides strong evidence that the estimated effects are robust to the inclusion of other predictors of child outcomes.

Finally, regarding children's cognitive and non-cognitive development, many studies have used valued-added model to deal with the potential omitted variable problem (Todd and Wolpin, 2003, 2007). This model includes the lagged outcome variables together with other control variables, assuming the lagged outcome variable are a sufficient statistic for all historical inputs and unobserved family and child characteristics that may affect children's outcome.[43] However, since a value-added model with the individual fixed effect would require observing the sample children for more than two periods, it is not applicable to our main estimation. Nevertheless, we estimated the value-added model without child fixed effect as a robustness check. We use the 2004 wave as our baseline sample, incorporating non-cognitive scores in wave 1 as commonly specified in the value-added model. Table 3.19 reports the estimates for a value-added model with and without instrumental variables. Our baseline results are robust to this alternative model specification.

[43]This method has been applied by Bono et al. (2016) and Fiorini and Keane (2014).

3.6 Mechanisms

Our results so far suggest a negative impact of the father's migration on child non-cognitive skill. It is then natural to ask how migration might have affected children's development. We test several different channels that are found to be salient for children's development broadly, including parental financial inputs, father's bargaining power, parental time inputs, child time inputs, parenting style, and parent health conditions. All the results are based on our instrumental variable model with child fixed effects.[44]

3.6.1 Parental Financial Inputs

As suggested in Hao and Yeung (2015) and Meyer and Sullivan (2008), household income is different from household consumption expenditure. The former measures the available financial resources and the latter measures the resource allocation. Distinguishing between the two will help clarify the mechanisms behind the effect of the father's migration. Thus, we explore not only how the father's migration affects household income but also how it affects children-specific spending.

In Table 3.10 panel A, we estimate a fixed effects instrumental variable model considering three outcome variables related to household income: total income, agriculture income, and labor income. As expected, households with a migrant father have significant lower agriculture income, but much higher labor income. Overall, the

[44]Results from alternative model specifications are available upon request.

migrant father significantly increased household income by 76 percent.

In panel B, we explore the effect of the father's migration on children-specific spending, which includes annual expenditure on children's clothes, school supplies, and children's allowance. Surprisingly, our results show that the father's migration significantly reduces household expenditure on children's clothes (67 percent) and has no significant impact on the expenditure of school supplies or allowance. These results imply that increased household income does not seem to be spent on improving children's living condition. One may argue that households with a migrant father have spent the increased income on education materials which is not covered in panel B. We are not able to pin down the exact amount of expenditure on studying materials because the GSCF provides no such information; however, we are able to test whether the father's migration has increased household expenditure on studying materials by exploring whether households have purchased certain study-related items for their children. In panel C, we investigate whether the household has purchased any reading material, a desk, or a dictionary for the children. Consistent with results in panel B, we observe that the father's migration is associated with a lower probability of the child having reading materials other than a test book or dictionary.[45]

Overall, results in Table 3.10 suggest that the father's migration increased household income, but it does not necessarily increase and actually decreases household material inputs for children development. Considering that material inputs are im-

[45]In contrast, we find that household with a migrating father spend more on transportation and adult clothing. These results are available upon request.

portant for children's non-cognitive development (Cunha and Heckman, 2008; Cunha et al., 2010), the fact that the father's migration reduces the material inputs is a potential mechanism behind its negative impact on children's non-cognitive skills. [46]

3.6.2 Parental Time Inputs

The father's migration can affect the father and mother's time allocation, which in turn would affect their time inputs on their children. On the one hand, the migrant father is absent for a substantial period of time, which reduces the time he can spend with his children. On the other hand, the mother who is left behind with the children may need to devote more time to other household matters, such as home production and farm work, when her spouse is away. It is intuitive that the father's migration likely reduces parental time inputs on children. However, we are reluctant to jump to the conclusion that the mother will reduce her time inputs on children because she might sacrifice her leisure time to compensate for the time lost caused by the father's migration.

To take a close look at this time input mechanism, we consider three variables that measure the father's and mother's time allocation, which include weekly hours of home production, weekly hours helping children with their homework, and playing or talking with children. For the migrant father, these variables reflect weekly hours

[46] It is worth noting that the father's migration significantly increased household spending on adult clothing and transportation, which explains where the additional income has been spent among families with a migrant father.

he spends on those activities when he stays at home. Thus, the estimates largely underestimated the negative impact of the father's migration on the father's time inputs on children. Nevertheless, as presented in panel A of Table 3.11, we find that the migrant father spends less time helping his children with their homework and less time playing and talking with his children even when the father is at home. Though the estimate on homework time is insignificant at -0.153, the effect on play and talk is significant at -0.500, which shows that they spend about 50 percent less time playing and talking with their children than fathers who stay home. Why do migrant fathers spend less time with their children when they come back? Shouldn't they spend more time to make up for the lost time when they are away? There are two potential explanations. First, a long-term absence makes it harder for the father to bond with his children and creates a communication barrier between the two. Second, the migrant father may have to devote substantial time to deal with family matters, such as repairing the house or socializing with other village members.[47]

In panel B of Table 3.11, we present how the father's migration affects the mother's time allocation considering the same set of activities, including home production, helping children with homework, and playing or talking with children. As expected, column (1) shows that a mother with a migrant spouse is much more likely to engage in home production activities, which more or less explains why she spends less

[47]There is also descriptive evidence suggesting that the longer a father migrates out, the less time input he spends with his children. However, migration duration itself is an endogenous choice, and is therefore beyond the scope of this paper. Another data restriction is the variables measuring father's migration duration in 2000 and 2004 have different definitions.

time communicating with her children as shown in column (3). Considering that the literature suggests investment in time plays an important role in determining children's non-cognitive development,[48] these results suggest that the lack of time inputs seems to be the main driving force behind the adverse effects of father's migration on children's non-cognitive skills.

3.6.3 Child Time Allocations

We also analyze the impact of the father's migration on time allocation made by children themselves. The GSCF asks the target child about his or her weekly hours spent on participating in home production, taking care of sick or elderly family members, doing homework, and watching TV. The estimation results are reported in the third panel of Table 3.11. Columns (1) and (2) show that the father's migration significantly increases the time a child spends on home production and caring for others by 67 and 71 percent, respectively. Columns 3 and 4 show that the father's migration has no significant effect on the time the child spends on doing homework or watching TV. These results are consistent with Chen (2013), which exams the time allocation of left-behind children in China and finds that they spend more time in household work. Antman (2011) also studies children's time allocation and finds that

[48]Since most socioeconomic surveys lack appropriate measures of parental time, most studies have used mother's employment to proxy it. Only until recently have a few papers started to use time-use diaries to measure parent time investments. See Del Boca et al. (2017, 2014); Fiorini and Keane (2014) for examples of papers that have used time-use diaries to study the effects of parental time inputs on child outcomes in Australia, the United Kingdom, and the United States.

children reduce study hours and increase work hours outside the home in response to a father's U.S. migration. The explanation of their results is that a father's migration is followed by a period of financial hardship for families in Mexico who may be financing the father's trip and also waiting for him to find a job in the United State. During this time period, children have to take financial responsibility for the household. Thus, they have to shift focus from schooling toward work outside of home. Our results are partially consistent with their findings, in the sense that we both find that left-behind children work more than other children. Our results are different from theirs in the sense that we find no significant impact on weekly hours doing homework. There are two reasons our results are different. First, their data measure total study hours, whereas we can only observe weekly hours of doing homework, which may be a major proportion of study hours. But it is possible that the father's migration decreases the total study hours by decreasing the time the child spends on reading or other studying activities. Second, it is possible that Chinese parents commonly have a high expectation about their children's education so they would make sure that their children spend enough time on studying before letting them engage in work-related activities. This explains why we observe nonsignificant effects of the father's migration on children's weekly hours of doing homework.

Spending more time on home production and caring for others may not have a direct negative impact on child development, but it implies that the child may have less time for other social activities that might be beneficial to their emotional and

social well-being. The GSCF does not have questions regarding weekly hours of other major activities besides what we listed here. Thus, our results only provide indirect evidence that the father's migration affects children's own time allocation. Moreover, the limited information from the GSCF makes it impossible to distinguish between the time the child spends on his or her own and the time actively supervised by a parent. Thus, the purpose of the above estimation is not to compare the importance of parental time and child time investment, but to provide insights on the possible mechanisms behind the negative effect of the father's migration. Future work testing this channel requires a time diary of the targeted children. such as such as the one in Child Development Supplement (CDS) of the Panel Survey of Income Dynamics (PSID) ,that can be used to estimate the effect of parental migration on the time allocation of children more comprehensively.

3.6.4 Father's Bargaining Power

The father's migration can affect not only how many financial resources are available to children by affecting household income but also how parents allocate these resources. Table 3.12 shows how the father's migration affects the decision power of the father relative to the mother regarding a wide range of family decisions. Columns (1)-(3) show that the father's migration increases the probability that he makes decisions on the children's schooling, on how to treat children, and on how to manage the family finances. Columns (4)-(6) show that the father's migration is positively related

with his decision power regarding the purchase of durable goods, planting crops, and transaction of livestock, though the estimated coefficients are not significant. These results are consistent with the literature, which finds that greater economic resources lead to greater bargaining power (Antman, 2014).

Overall, results in Table 3.12 suggest that the father's migration increases his bargaining power, especially related to children's education and financial inputs. The existing studies show that fathers, compared with mothers, are less likely to spend family resources on their children (Hoddinott and Haddad, 1995; Lépine and Strobl, 2013; Li and Wu, 2011). Our results suggest that fathers' migration may reduce the amount of resources actually allocated to their children, which to some extent, explains why households with migrant fathers spend less on children's clothes and are less likely to purchase study materials given that their household income has been increased (Table 3.10).

3.6.5 Parenting Style and Parental Health

In addition to the mechanisms we have tested so far, the literature has emphasized other important determinants of child development. One is parenting style (Doepke and Zilibotti, 2017; Dooley and Stewart, 2007; Fiorini and Keane, 2014). Although there are no consistent measurements of parental style so far, the few papers that considered these kinds of variables commonly find that parenting style has an impact on child developmental outcomes. For example, Dooley and Stewart (2007) find that

different aspects of parenting (positiveness, hostility, consistency, and punitiveness) are more important than family income in terms of determining children's behavior and emotional outcomes. Fiorini and Keane (2014) find that parental warmth and effective discipline lead to better non-cognitive outcomes for children. More recently, Doepke and Zilibotti (2017) develops a theory to explain how parenting style plays an important role in determining children's welfare and economic success. Parenting style can be an important channel through which parental migration affects child development. The GSCS provides limited information on parenting style. Nevertheless, we define harsh parenting style by the mother's and child's answer to the question "whether you beat your child/you are beaten when misbehaved." Using the baseline fixed effects model with instrumental variables, our results in Appendix Table 3.20 show that the effect of the father's migration on parenting style is insignificant. These results, however, do not imply that parenting style is not important in terms of explaining the link between a father's migration and his children's development outcomes. Future work needs to measure parenting styles in a more comprehensive way to better test this mechanism.

The other determinant that was mentioned in the literature is parent's health condition, which may have an impact on the quality of time parents spend with their children. For example, Ronda (2016) and Herbst (2017) find that maternal psychological distress and depression may have a negative impact on children's outcomes. In the meantime, the migration literature has provided some evidence that migra-

tion increases the probability of being in poor physical and mental health (Antman, 2010; Barrett and Mosca, 2013). To test this mechanism, we investigate whether the father's migration has an impact on the self-assessed health status of the father and mother.[49] Appendix Table 3.20 provides the estimation results, showing that the father's migration does not increase the probability of parents in poor health condition. However, we are reluctant to conclude that parent health status is not an important mechanism behind the link between parental migration and a child's development, considering that our measure of health status is self-reported and does not differentiate between mental health and physical health. This is a mechanism worth testing in the future. [50]

3.7 Discussion of Sub-sample Results

Table 3.13 reports the sub-sample estimation results and highlights some interesting patterns.[51] First, columns (1) and (2) show that the effects of father's migration are larger among boys (-0.057 and -0.065) than those for girls (-0.028 and -0.047). We hypothesize that there are several explanations for this gender difference. First,

[49]The assessment of the father's and mother's health condition is classified into five categories: very poor, poor, average, good, and very good. We group the first two and generate a dummy variable to indicate whether the father or mother has poor health.

[50]Parents make decisions taking as given the production function of human capital. It is possible that parental migration will change parents view on the production function regarding the expectation of the economic return of their investment. However, without valid measurements on these variables, it is difficult to test these hypotheses empirically using our data.

[51]As is often the case, however, the instruments are much weaker by subgroup, and the F statistic on the excluded instruments is only above 10 for the children whose father did not graduate from elementary school. The results should thus be interpreted with caution.

the left-behind mother communicates better with girls while having difficulty dealing with teen boys. Second, lacking a positive male role model has a negative impact on a boy's development. Thus, maternal input could be a poorer substitute for the lack of paternal inputs when it comes to raising a boy. Third, as proposed in Bertrand and Pan (2013), the non-cognitive development of boys, unlike that of girls, appears extremely responsive to the quality of parental inputs, which is negatively associated with parental migration.[52] Nonetheless, we are not intending to decompose these different channels, and we leave it for future studies. This kind of gender difference has also been documented in recent child development literature, which finds that father absence, usually caused by divorce, has a greater influence on boys than girls. For example, Bertrand and Pan (2013) find that boys do especially poorly and are much more likely to develop behavior problems in broken families, which are usually associated with worse parental inputs. Figlio et al. (2019) find that family disadvantage disproportionately impedes the development of boys by using birth certificates matched to schooling records in Florida.

Columns (3)-(6) divide the sample children by their parents' education levels. We find that the negative effect of the father's migration is more evident if parents, especially mothers, are less educated. For example, if the mother has not graduated from elementary school, then the father's migration reduces internalizing and externalizing scores by 0.036 and 0.071 points. But if the mother has graduated from elementary

[52] Antman (2012) finds that a left-behind mother allocate more resources on girls relative to boys while the father migrates to the United States. However, this explanation may not apply in the context of rural China, considering the preference of boys over girls.

school, the two coefficients are reduced to 0.025 and 0.05. These results suggest that mothers with better education are more likely to engage in high-quality parenting and can better mitigate the adverse effect caused by the father's absence (Carneiro et al., 2013).

Moreover, as previously discussed, lack of time inputs seems to be the major mechanism through which the father's migration affects his children's development. In that sense, it is possible that better infrastructure, such as having access to a telephone service, may help left-behind children connect and communicate with their fathers. In columns (7) and (8) of Table 3.13, we divide our sample children by whether their resident village has telephone services, but the negative effect of father's migration does not vary much across the two groups. This may be because telephone communication is not a good substitute for face-to-face communication in terms of parenting or because the telephone is not widely used among migrant families to maintain effective communication between the migrant father and his children left behind.

3.8 Conclusion

This paper sets out to identify the effects of a father's migration on his children's non-cognitive development. By exploiting a longitudinal data set and using instrumental variables based on the destination provinces, we identify a negative effect of

a father's migration and effects are stronger for boys than girls. We explore several possible mechanisms behind this negative effect. First, we find that a fathers' migration is associated with an increase in family income, which, however, does not lead to an increase in child-related spending. In the meanwhile, our results show that a father's migration increases his decision power on financial allocation within the family, which provides a potential explanation of the observed divergence in family income and spending. More important, a father's migration reduces the time both parents spend talking and playing with their children. Overall, the father's migration reduces both parent financial and time inputs, which might be the two major driving forces behind the negative effect of a father's migration on left-behind children.

These results expand the scope of current literature by uncovering negative consequences of parental migration that is rarely studied in the literature. More important, our discussion on potential mechanisms suggest that relevant policies are urgently needed to help migrant parents improve their financial and time inputs on children. Considering that non-cognitive skill is a vital dimension of human capital, these policies would have an important impact on increasing inter-generational mobility and on reducing rural-urban inequality.

3.9 Figures and Tables

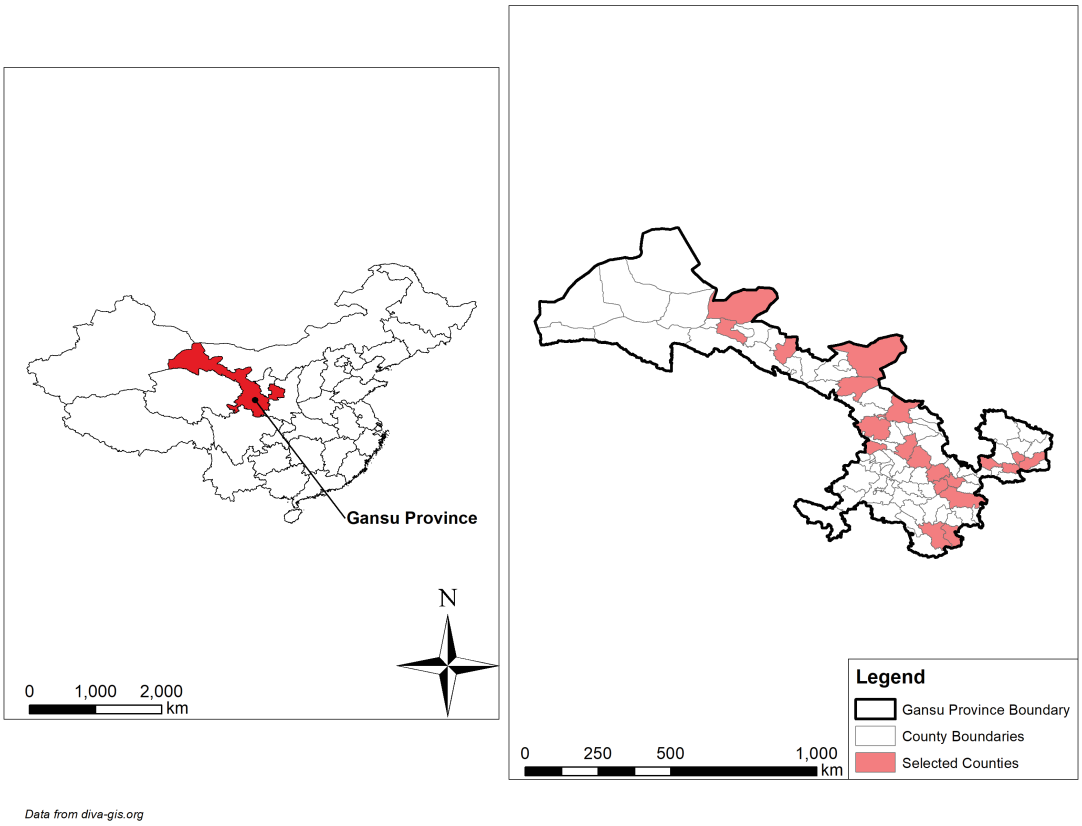


Figure 3.1: A Map of Gansu Province in China

Table 3.1: Father's Migration Status by Year

Panel A	2000		2004		Total	
	No.	%	No.	%	No.	%
Father Non-Migrant	1,119	62.9%	1,159	65.1%	2,278	64.0%
Intra-provincial Migrant	548	30.8%	471	26.5%	1,019	28.6%
Inter-provincial Migrant	112	6.3%	149	8.4%	261	7.3%
Total	1,779	100.0%	1,779	100.0%	3,558	100.0%
Panel B Inter-provincial Migrant	No.	%				
only in 2000	66	3.7%				
only in 2004	93	5.2%				
in both years	46	2.6%				
in neither year	1574	88.5%				
Total	1,779	100.0%				

Data source: Gansu Survey of Children and Families (GSCF) 2000 and 2004 waves.

Note: A father's migration status is defined by his workplace location. In particular, we define a father to be a non-migrant if his workplace is located in the village, an intra-provincial migrant if his workplace is located outside of the village but within Gansu province and an inter-provincial migrant if his workplace is located in another province outside of Gansu.

Table 3.2: Summary Statistics by Father's Migration Status

	Total	Father Non-migrant	Intra-provincial Migrant	Inter-provincial Migrant
Internalizing Score	0.000	0.019	-0.014	-0.115
	(1.000)	(0.979)	(1.027)	(1.063)
Externalizing Score	0.000	0.003	0.013	-0.078
	(1.000)	(0.986)	(1.036)	(0.984)
Child Female	0.469	0.464	0.485	0.441
	(0.499)	(0.499)	(0.500)	(0.497)
Child's Age	12.999	13.019	12.892	13.240
	(2.281)	(2.286)	(2.262)	(2.286)
Father's Age	39.493	39.865	38.919	38.485
	(5.512)	(5.821)	(4.732)	(5.261)
Mother's Age	37.113	37.426	36.636	36.240
	(4.958)	(5.221)	(4.200)	(5.091)
Father's Years of Eduction	6.757	6.648	7.213	5.931
	(4.154)	(4.164)	(4.050)	(4.283)
Mother's Years of Eduction	3.749	3.761	4.066	2.402
	(4.014)	(4.005)	(4.089)	(3.499)
Number of Siblings	1.329	1.338	1.289	1.406
	(0.719)	(0.733)	(0.690)	(0.699)
Grandparent at Home	0.206	0.206	0.204	0.218
	(0.405)	(0.404)	(0.403)	(0.414)
Village Population (1000)	1.712	1.677	1.872	1.384
	(1.809)	(1.549)	(2.431)	(0.653)
Village has Primary School	0.830	0.806	0.871	0.881
	(0.375)	(0.396)	(0.335)	(0.324)
Village has Middle School	0.101	0.104	0.099	0.077
	(0.301)	(0.306)	(0.299)	(0.267)
Dest. Wage	12095.361	12159.194	11818.068	12620.841
	(4720.891)	(4820.492)	(4526.407)	(4534.633)
Dest. Employment (1000)	13331.346	13620.136	13307.845	10902.550
	(12220.901)	(12167.200)	(12332.639)	(12023.828)
Dest. Wage in Manuf.	10637.723	10685.031	10423.401	11061.584
	(3554.708)	(3608.145)	(3414.807)	(3580.332)
Dest. Employment in Manuf. (1000)	774.203	787.806	765.992	687.539
	(845.585)	(841.017)	(860.453)	(824.165)
Dest. Wage in Constr.	15332.017	15360.850	15163.207	15739.431
	(4413.732)	(4485.089)	(4315.598)	(4139.841)
Dest. Employment in Constr. (1000)	234.966	237.391	232.944	221.700
	(183.421)	(180.854)	(185.469)	(197.204)
Observations	3558	2278	1019	261
% of Total Sample	100%	64.0%	28.6%	7.4%

Data source: Gansu Survey of Children and Families (GSCF) 2000 and 2004 waves. Note: The sample includes children who appear in both 2000 and 2004 waves. Internalizing and externalizing scores have been standardized to have means equal to 0 and standard deviations equal to 1. Higher internalizing or externalizing scores indicate higher non-cognitive skills. Monetary value in 2004 real RMB.

Table 3.3: Migration Destination Provinces

	2000		2004		Total	
	No.	%	No.	%	No.	%
Xinjiang	670	37.7%	693	39.0%	1,363	38.3%
Guangdong	225	12.6%	262	14.7%	487	13.7%
Ningxia	147	8.3%	147	8.3%	294	8.3%
Beijing	92	5.2%	110	6.2%	202	5.7%
Qinghai	89	5.0%	108	6.1%	197	5.5%
Shannxi	72	4.0%	55	3.1%	127	3.6%
Inner Mongolia	37	2.1%	53	3.0%	90	2.5%
Zhejiang	19	1.1%	50	2.8%	69	1.9%
Shanxi	45	2.5%	15	0.8%	60	1.7%
Tibet	18	1.0%	36	2.0%	54	1.5%
Fujian	18	1.0%	18	1.0%	36	1.0%
Tianjing	0	0.0%	33	1.9%	33	0.9%
Shanghai	0	0.0%	18	1.0%	18	0.5%
Gansu	347	19.5%	181	10.2%	528	14.8%
Total	1,779	100.0%	1,779	100.0%	3,558	100.0%

Data source: Gansu Survey of Children and Families (GSCF) 2000 and 2004 waves.

Note: Migration destination is a village-level variable. For example, in 2000, 670 observations in our sample live in villages where Xinjiang is reported as the most popular destination other than Gansu; 347 observations live in villages where no provinces other than Gansu are reported as migration destination.

Table 3.4: Parental Financial inputs by Father's Migration Status

	Total	Non-migrant	Intra-provincial Migrant	Inter-provincial Migrant
Annual Income (RMB)				
Agriculture	6311.93	7015.41	5387.53	3781.05
	(6231.03)	(6705.68)	(5371.45)	(3206.67)
Labor	2721.10	1412.29	4854.39	5815.44
	(3655.96)	(2920.97)	(3518.95)	(4146.84)
Total	10121.1	10186.3	10040.4	9867.8
	(10048.1)	(11005.4)	(7813.8)	(9021.3)
Annual Expenditure (RMB)				
Children Clothes	234.8	244.1	228.6	178.3
	(178.7)	(184.2)	(171.2)	(143.9)
School Supplies	35.4	36.8	33.8	29.8
	(40.4)	(42.3)	(37.3)	(34.1)
Children's Allowance	93.4	98.1	85.9	82.3
	(144.2)	(149.8)	(134.3)	(129.8)
Item (Dummy)				
Reading Materials	0.604	0.623	0.612	0.413
	(0.489)	(0.485)	(0.488)	(0.493)
Dictionary	0.956	0.952	0.968	0.945
	(0.204)	(0.213)	(0.175)	(0.229)
Desk or Bookshelf	0.626	0.630	0.645	0.520
	(0.484)	(0.483)	(0.479)	(0.501)

Data source: Gansu Survey of Children and Families (GSCF) 2000 and 2004 waves.

Note: Total income includes agriculture and labor income. All monetary variables were converted to real terms (the year 2000 RMB value) using the national Consumer Price Index (CPI).

Table 3.5: Parental and Children's Own Time Inputs by Father's Migration Status

	Total	Non-migrant	Intra-provincial Migrant	Inter-provincial Migrant
Father Weekly Hours				
Home Production	9.368	9.795	7.727	12.178
	(11.559)	(11.500)	(10.266)	(16.166)
Help Children with Homework	1.769	1.892	1.592	1.185
	(3.062)	(3.165)	(2.943)	(2.183)
Play or Talk with Children	2.883	2.958	2.790	2.454
	(3.741)	(3.826)	(3.613)	(3.290)
Mother Weekly Hours				
Home Production	30.172	29.125	31.539	34.096
	(18.936)	(14.042)	(26.970)	(17.158)
Help Children with Homework	1.319	1.264	1.557	0.843
	(2.688)	(2.607)	(2.982)	(1.964)
Play or Talk with Children	3.773	3.668	4.038	3.650
	(4.928)	(4.988)	(4.975)	(4.099)
Child Weekly Hours				
Home Production	3.20	2.95	3.44	4.46
	(5.44)	(5.17)	(5.71)	(6.36)
Take Care Others	0.59	0.55	0.61	0.89
	(1.91)	(1.87)	(1.86)	(2.33)
Do Homework	10.05	9.78	10.50	10.63
	(6.51)	(6.38)	(6.56)	(7.34)
Watch TV	6.32	6.26	6.42	6.35
	(5.05)	(5.01)	(4.99)	(5.59)

Data source: Gansu Survey of Children and Families (GSCF) 2000 and 2004 waves.

Note: For the migrant father, weekly hours variables reflect the time he spends when he stays at home.

Table 3.6: Father's Bargaining Power by Father's Migration Status

	Total	Non-migrant	Intra-provincial Migrant	Inter-provincial Migrant
Decision on Child's School	0.18	0.20	0.15	0.19
	(0.39)	(0.40)	(0.36)	(0.39)
Decision on How to Treat the Child	0.23	0.24	0.22	0.24
	(0.42)	(0.43)	(0.41)	(0.43)
Decision on Purchase of New Durable Goods	0.23	0.24	0.20	0.22
	(0.42)	(0.43)	(0.40)	(0.42)
Decision on Management of Family Finance	0.42	0.44	0.38	0.41
	(0.49)	(0.50)	(0.49)	(0.49)
Decision on Planting of Crops	0.30	0.33	0.27	0.24
	(0.46)	(0.47)	(0.44)	(0.43)
Decision on Livestock Transaction	0.34	0.37	0.31	0.30
	(0.48)	(0.48)	(0.46)	(0.46)

Data source: Gansu Survey of Children and Families (GSCF) 2000 and 2004 waves. Note: Questions regarding the power of decision-making in the family are reported by the mother. If the father makes decisions, the variables are coded as 1; if the mother makes the decision or the parents make decision together, the variables are coded as 0.

Table 3.7: Father's Migration Status and Child Non-cognitive Skills: Main Results

	Internalizing Score				Externalizing Score			
	OLS	2SLS	FE	FE+IV	OLS	2SLS	FE	FE+IV
Father Migration	-0.046***	-0.058**	-0.025***	-0.029*	-0.084***	-0.088***	-0.054***	-0.058***
	(0.019)	(0.031)	(0.009)	(0.014)	(0.018)	(0.030)	(0.012)	(0.016)
F test		17.24		35.82		17.24		35.82
Over-identification test		0.11		0.79		0.11		0.79
Observations	3558	3558	3558	3558	3558	3558	3558	3558
R^2	0.205	0.202	0.013	0.008	0.029	0.028	0.014	0.009

Note: Standard errors in brackets are significant at *10% **5%, and ***1% and clustered at the county level. The dependent variables are the internalizing and externalizing scores. Higher internalizing or externalizing scores indicate higher non-cognitive skills. The instrumental variables include the employment and wage in the top destination provinces provided in the China Labor Statistical Yearbook. For OLS and 2SLS specifications, control variables include child's age and gender, parental age and education, number of siblings, whether any grandparent resides at home or not, village level population, whether has elementary school, and whether has middle school, as well as county and year fixed effects. For FE and FE with IV specifications, time-invariant variables such as child's gender, parental education, number of siblings, county fixed effect are excluded. Parental ages are also excluded from FE and FE with IV specifications, due to co-linearity with the variation of child's age.

Table 3.8: Father's Migration Status and Child Non-cognitive Skills: First-Stage Regressions

	Total IV		Manufacturing IV		Construction IV	
	2SLS	FE	2SLS	FE	2SLS	FE
Dest. Employment (1000)	0.002*	0.001				
	(0.001)	(0.001)				
Dest. Wage (1000)	0.014***	0.010***				
	(0.004)	(0.002)				
Dest. Employment in Manuf. (1000)			0.005***	0.005***		
			(0.001)	(0.001)		
Dest. Wage in Manuf. (1000)			0.021***	0.014***		
			(0.005)	(0.003)		
Dest. Employment in Constr. (1000)					0.003***	0.006***
					(0.001)	(0.001)
Dest. Wage in Constr. (1000)					0.010***	0.011***
					(0.002)	(0.002)
Observations	3558	3558	3558	3558	3558	3558
R^2	0.125	0.040	0.137	0.045	0.143	0.049

Note: Standard errors in brackets are significant at *10% **5%, and ***1% and clustered at the county level. The dependent variables are the internalizing and externalizing scores. Higher internalizing or externalizing scores indicate higher non-cognitive skills. The instrumental variables include the employment and wage in the top destination provinces provided in the China Labor Statistical Yearbook. For 2SLS specifications, control variables include child's age and gender, parental age and education, number of siblings, whether any grandparent resides at home or not, village level population, whether has elementary school, and whether has middle school, as well as county and year fixed effects. For FE with IV specifications, time-invariant variables such as child's gender, parental education, number of siblings, county fixed effect are excluded. Parental ages are also excluded from FE and FE with IV specifications, due to co-linearity with the variation of child's age.

Table 3.9: Father's Migration Status and Child Non-cognitive Skills: Alternative Instrumental Variables

	Internalizing Score				Externalizing Score			
	Manufacturing IV		Construction IV		Manufacturing IV		Construction IV	
	2SLS	FE+IV	2SLS	FE+IV	2SLS	FE+IV	2SLS	FE+IV
Father Migration	-0.054***	-0.034***	-0.044*	-0.024	-0.084***	-0.056***	-0.087***	-0.068*
	(0.024)	(0.017)	(0.022)	(0.018)	(0.033)	(0.023)	(0.029)	(0.032)
F test	19.28	37.87	15.31	34.64	19.28	37.87	15.31	34.64
Over-identification test	0.09	0.11	0.91	0.77	0.09	0.11	0.91	0.77
Observations	3558	3558	3558	3558	3558	3558	3558	3558
R^2	0.010	0.011	0.005	0.004	0.020	0.020	0.006	0.005

Note: Standard errors in brackets are significant at *10% **5%, and ***1% and clustered at the county level. The dependent variables are the internalizing and externalizing scores. Higher internalizing or externalizing scores indicate higher non-cognitive skills. The instrumental variables include the employment and wage in the top destination provinces provided in the China Labor Statistical Yearbook. For 2SLS specifications, control variables include child's age and gender, parental age and education, number of siblings, whether any grandparent resides at home or not, village level population, whether has elementary school, and whether has middle school, as well as county and year fixed effects. For FE with IV specifications, time-invariant variables such as child's gender, parental education, number of siblings, county fixed effect are excluded. Parental ages are also excluded from FE and FE with IV specifications, due to co-linearity with the variation of child's age.

Table 3.10: Mechanisms 1: Parental Financial Inputs

Panel A	log(Income)		
	Total	Agriculture	Labor
Father Migration	0.761***	-0.929***	11.944***
	(0.215)	(0.243)	(1.110)
F test	35.82	35.82	35.82
Over-identification test	1.85	2.97	1.16
Observations	3558	3558	3558
RMSE	0.705	0.533	4.602
Panel B	log(Expenditure)		
	Children Clothes	School Supplies	Allowance
Father Migration	-0.671***	-0.291	0.055
	(0.249)	(0.273)	(0.249)
F test	35.82	35.82	35.82
Over-identification test	2.01	0.06	1.84
Observations	3558	3558	3558
RMSE	0.770	0.997	0.892
Panel C	Items (dummy)		
	Reading Materials	Desk	Dictionary
Father Migration	-0.378***	0.008	-0.254*
	(0.130)	(0.055)	(0.130)
F test	36.12	36.25	36.25
Over-identification test	3.99	0.30	0.53
Observations	3464	3494	3494
RMSE	0.422	0.192	0.421

Note: Total income includes farming, livestock, and labor income. All monetary variables were converted to real terms (the year 2000 RMB value) using CPI. Standard errors in brackets are significant at *10% **5%, and ***1% and clustered at the county level. The results are based on FE with IV specification used in column (4) and (8) in Table 3.7.

Table 3.11: Mechanisms 3: Parental and Children' Own Time Inputs

Panel A	Father Time Allocation			
	Home Production	Time with Children	Help Homework	Play and Talk
Father Migration	0.408	-0.540**	-0.153	-0.500**
	(0.320)	(0.246)	(0.215)	(0.230)
F test	51.01	51.01	51.05	51.07
Over-identification test	2.46	1.68	5.63	8.66
Observations	3224	3224	3226	3228
RMSE	1.021	0.864	0.740	0.830
Panel B	Mother Time Allocation			
	Home Production	Time with Children	Help Homework	Play and Talk
Father Migration	0.440***	-0.242	-0.022	-0.570**
	(0.167)	(0.246)	(0.196)	(0.240)
F test	36.13	36.10	36.10	36.19
Over-identification test	4.15	8.22	0.01	11.71
Observations	3490	3490	3490	3496
RMSE	0.565	0.865	0.672	0.851
Panel C	Child Time Allocation			
	Home Production	Take Care Others	Do Homework	Watch TV
Father Migration	0.676***	0.713**	-0.078	0.208
	(0.142)	(0.338)	(0.466)	(0.464)
F test	35.84	35.84	35.84	35.84
Over-identification test	0.27	0.07	0.43	0.01
Observations	3558	3558	3558	3558
RMSE	0.900	0.560	0.773	0.769

Note: Time variables are measured as weekly hours in log. Standard errors in brackets are significant at *10% **5%, and ***1% and clustered at the county level. The results are based on FE with IV specification used in column (4) and (8) in Table 3.7.

Table 3.12: Mechanisms 2: Father's Bargaining Power

	Child School	Treat Child	Family Finance	Purchase Durables	Plant Crops	Livestock Transaction
Father Migration	0.536**	0.525**	0.597**	0.215	0.094	0.228
	(0.255)	(0.229)	(0.251)	(0.277)	(0.257)	(0.270)
F test	36.55	36.69	35.57	37.22	36.95	36.58
Over-identification test	0.88	1.29	1.16	2.27	0.92	2.26
Observations	3474	3474	3474	3474	3478	3460
RMSE	0.424	0.380	0.410	0.462	0.429	0.448

Note: Questions regarding the power of decision-making in the family are reported by the mother. If the father makes decisions, the variables are coded as 1; if the mother makes the decision or the parents make decision together, the variables are coded as 0. Standard errors in brackets are significant at *10% **5%, and ***1% and clustered at the county level. The results are based on FE with IV specification used in column (4) and (8) in Table 3.7.

Table 3.13: Father's Migration Status and Child Non-cognitive Skills: Subsample Results

	Child Gender		Mother Education		Father Education		Village Phone Service	
	Boys	Girls	Low	High	Low	High	No	Yes
Internalizing Score								
Father Migration	-0.057***	-0.028*	-0.036*	-0.025	-0.033*	-0.028*	-0.030	-0.029
	(0.018)	(0.016)	(0.021)	(0.019)	(0.018)	(0.016)	(0.030)	(0.022)
F test	17.78	16.38	21.60	15.6	12.33	17.34	19.29	30.44
Over-identification test	0.08	0.00	0.31	0.03	0.49	0.27	0.91	0.08
Observations	1900	1658	1796	1762	806	2752	1502	2056
Externalizing Score								
Father Migration	-0.065***	-0.047***	-0.071***	-0.050***	-0.066***	-0.056***	-0.061***	-0.059***
	(0.022)	(0.017)	(0.016)	(0.023)	(0.019)	(0.026)	(0.018)	(0.025)
F test	17.78	16.38	21.60	15.6	12.33	17.34	19.29	30.44
Over-identification test	0.06	0.09	0.07	0.91	1.21	1.29	0.063	0.26
Observations	1900	1658	1796	1762	806	2752	1502	2056

Note: Low-educated mothers and fathers are those who did not graduate from the elementary schools. Standard errors in brackets are significant at *10% **5%, and ***1% and clustered at the county level. The results are based on FE with IV specification used in column (4) and (8) in Table 3.7.

3.10 Appendix

Table 3.14: Survey Questions on Children's Non-cognitive Skills in GSCF 2000 and 2004 waves

Externalization Behavioral Problems
◇ I break things on purpose
◇ I lose my temper with others
◇ Even if wrong, I am reluctant to listen to others
◇ I steal things from others or my home
◇ I like to show off my strengths in front of others
◇ I always want to be the center of attention
◇ I often quarrel with others
◇ I do not observe school discipline
◇ I like to brag
◇ It bothers me if others do things better than I do
◇ I act impulsively
◇ I often say obscenities
◇ I often make fun of others
◇ I sometimes tell lies
◇ I am easily angered
◇ I often disregard other people's ideas
◇ I sometimes menace and even hurt others
Internalization Behavioral Problems
◇ I don't want others to meddle in my own business
◇ I can't concentrate on what I am doing
◇ I have many strange/weird ideas (often daydream)
◇ I easily get flushed/frustrated/anxious
◇ I can't do things well when my parents are not present
◇ I am very indifferent to others
◇ I am very shy
◇ I am often teased by classmates
◇ I do not feel guilty, even if I have done something wrong
◇ I feel inferior to others
◇ I often am suspicious of others
◇ I prefer to be alone
◇ I often feel nervous
◇ I am often bored
◇ I stay quiet when I am with my classmates or friends
◇ There is always something to worry about

Table 3.15: Father's Migration Duration: by Year

	2000		2004			
	Months Away		Months Away Consecutively		Months Away During Weekdays	
	Mean	S.D.	Mean	S.D.	Mean	S.D.
Father Non-Migrant	0.428	(1.691)	0.082	(0.784)	0.040	(0.540)
Intra-provincial Migrant	3.09	(3.845)	2.014	(3.183)	0.679	(2.121)
Inter-provincial Migrant	6.109	(3.347)	4.466	(3.446)	0.912	(2.199)
Total	1.877	(3.289)	0.960	(2.428)	0.282	(1.377)

Note: A father's migration status is defined by his workplace location. In particular, we define a father to be a non-migrant if his workplace is located in the village, an intra-provincial migrant if his workplace is located outside of the village but within Gansu province and an inter-provincial migrant if his workplace is located in another province outside of Gansu.

Table 3.16: Summary Statistics by Father's Migration Status (2000)

	Total	Father Non-migrant	Intra-provincial Migrant	Inter-provincial Migrant
Internalizing Score	-0.003	0.026	-0.037	-0.138
	(1.001)	(0.960)	(1.050)	(1.146)
Externalizing Score	-0.003	0.010	-0.013	-0.078
	(1.001)	(0.968)	(1.051)	(1.079)
Child Female	0.469	0.467	0.487	0.393
	(0.499)	(0.499)	(0.500)	(0.491)
Child's Age	10.999	10.975	11.058	10.953
	(1.096)	(1.089)	(1.106)	(1.114)
Father's Years of Edu.	6.757	6.638	7.182	5.866
	(4.154)	(4.148)	(4.074)	(4.418)
Mother's Years of Edu.	3.749	3.777	4.057	1.955
	(4.014)	(4.001)	(4.103)	(3.189)
Father's Age	37.500	37.764	37.179	36.426
	(5.139)	(5.397)	(4.486)	(5.313)
Mother's Age	35.119	35.373	34.838	33.963
	(4.540)	(4.854)	(3.732)	(4.672)
Number of Siblings	1.329	1.332	1.297	1.464
	(0.717)	(0.715)	(0.714)	(0.734)
Grandparent at Home	0.227	0.238	0.199	0.250
	(0.419)	(0.426)	(0.400)	(0.435)
Village Population	1.594	1.583	1.631	1.515
	(0.819)	(0.832)	(0.808)	(0.732)
Village has Elem. School	0.713	0.666	0.801	0.750
	(0.453)	(0.472)	(0.400)	(0.435)
Village has Mid. School	0.049	0.045	0.062	0.027
	(0.216)	(0.207)	(0.241)	(0.162)
Dest. Wage	8763.128	8713.006	8774.322	9209.121
	(2248.299)	(2221.963)	(2187.512)	(2726.506)
Dest. Employment (1000)	12769.524	12854.951	12846.605	11538.878
	(10689.137)	(10429.568)	(11028.795)	(11556.224)
Dest. Wage in Manuf.	8120.001	8083.147	8099.077	8590.593
	(2047.102)	(2003.453)	(2022.871)	(2508.707)
Dest. Employment in Manuf. (1000)	777.959	775.408	778.677	799.938
	(773.425)	(752.567)	(804.270)	(830.174)
Dest. Wage in Constr.	12218.446	12142.979	12292.206	12611.541
	(2444.164)	(2404.783)	(2476.840)	(2640.515)
Dest. Employment in Constr. (1000)	231.081	230.957	227.297	250.830
	(173.496)	(168.574)	(177.901)	(198.764)
Observations	1779	1159	471	149
% of Total Sample	100%	65.1%	26.5%	8.4%

Data source: Gansu Survey of Children and Families (GSCF) 2000 waves. Note: The sample includes children who appear in both 2000 waves. Higher internalizing or externalizing scores indicate higher non-cognitive skills. Monetary value in 2004 real RMB.

Table 3.17: Summary Statistics by Father's Migration Status (2004)

	Total	Father Non-migrant	Intra-provincial Migrant	Inter-provincial Migrant
Internalizing Score	0.003	0.012	0.013	-0.097
	(0.999)	(0.998)	(1.001)	(0.999)
Externalizing Score	0.003	-0.003	0.043	-0.077
	(0.999)	(1.003)	(1.018)	(0.910)
Child Female	0.468	0.462	0.482	0.477
	(0.499)	(0.499)	(0.500)	(0.501)
Child's Age	14.999	14.993	15.026	14.960
	(1.096)	(1.094)	(1.084)	(1.146)
Father's Years of Edu.	6.757	6.657	7.248	5.980
	(4.154)	(4.182)	(4.026)	(4.192)
Mother's Years of Edu.	3.749	3.745	4.076	2.738
	(4.014)	(4.009)	(4.078)	(3.690)
Father's Age	41.486	41.894	40.943	40.034
	(5.140)	(5.489)	(4.177)	(4.672)
Mother's Age	39.106	39.408	38.729	37.951
	(4.540)	(4.781)	(3.721)	(4.721)
Number of Siblings	1.329	1.345	1.278	1.362
	(0.722)	(0.750)	(0.663)	(0.670)
Grandparent at Home	0.186	0.175	0.210	0.195
	(0.389)	(0.380)	(0.408)	(0.397)
Village Population	1.829	1.768	2.152	1.286
	(2.419)	(2.007)	(3.449)	(0.570)
Village has Elem. School	0.948	0.941	0.953	0.980
	(0.223)	(0.235)	(0.211)	(0.141)
Village has Mid. School	0.153	0.162	0.142	0.114
	(0.360)	(0.369)	(0.350)	(0.319)
Dest. Wage	15427.595	15486.445	15359.412	15185.356
	(4160.679)	(4285.874)	(3929.130)	(3888.306)
Dest. Employment (1000)	13893.168	14358.912	13844.490	10424.236
	(13560.810)	(13600.187)	(13686.885)	(12380.715)
Dest. Wage in Manuf.	13155.445	13197.117	13127.709	12918.973
	(2898.862)	(2977.575)	(2620.444)	(3115.359)
Dest. Employment in Manuf. (1000)	770.447	799.776	751.233	603.051
	(912.259)	(918.538)	(922.172)	(812.180)
Dest. Wage in Constr.	18445.588	18467.664	18503.565	18090.597
	(3687.753)	(3782.162)	(3522.433)	(3450.991)
Dest. Employment in Constr. (1000)	238.851	243.602	239.513	199.804
	(192.807)	(191.835)	(193.886)	(193.814)
Observations	1779	1119	548	112
% of Total Sample	100%	62.9%	30.8%	6.3%

Data source: Gansu Survey of Children and Families (GSCF) 2004 waves. Note: The sample includes children who appear in 2004 waves. Higher internalizing or externalizing scores indicate higher non-cognitive skills. Monetary value in 2004 real RMB.

Table 3.18: Father's Migration Status and Child Non-cognitive Skills: Additional Controls

	Internalizing Score					Externalizing Score				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Father Migration	-0.027*	-0.029*	-0.021	-0.030*	-0.024*	-0.036***	-0.051***	-0.044***	-0.054***	-0.057***
	(0.014)	(0.015)	(0.014)	(0.015)	(0.014)	(0.016)	(0.017)	(0.016)	(0.018)	(0.019)
Child Poor Health		-0.049					-0.068			
		(0.093)					(0.085)			
Chinese Test Score		0.043					0.057			
		(0.039)					(0.035)			
Student Violate Discipline			-0.082**					-0.030		
			(0.041)					(0.040)		
Student Disrupt the Class			-0.048					0.003		
			(0.043)					(0.041)		
Teacher Absent			-0.223***					-0.146***		
			(0.047)					(0.045)		
School Close			-0.264***					-0.324***		
			(0.044)					(0.043)		
Father Poor Health				0.075					-0.043	
				(0.072)					(0.069)	
Mother Poor Health				-0.096					-0.009	
				(0.069)					(0.066)	
Log(Household Income)					0.007					0.020
					(0.036)					(0.035)
Father log (Hours with Children)					0.093**					0.062**
					(0.041)					(0.030)
Mother log (Hours with Children)					0.051					0.045
					(0.041)					(0.039)
Constant	2.753	1.830	1.959	-0.341	1.205	2.657	2.349	0.605	-1.975	-0.322
	(4.516)	(5.996)	(4.205)	(4.638)	(4.643)	(4.233)	(5.473)	(4.058)	(4.438)	(4.406)
County × Year dummies	Yes	No	No	No	No	Yes	No	No	No	No
F test	35.82	27.78	33.22	34.67	34.89	35.82	27.78	33.22	34.67	34.89
Over-identification test	0.72	0.89	0.79	0.77	0.77	0.72	0.89	0.79	0.77	0.77
Observations	3558	2547	3508	3529	3178	3558	2547	3508	3529	3178
R^2	0.007	0.011	0.008	0.008	0.009	0.007	0.011	0.008	0.008	0.009

Note: Standard errors in brackets are significant at *10% **5%, and ***1% and clustered at the county level. The dependent variables are the internalizing and externalizing scores. Higher internalizing or externalizing scores indicate higher non-cognitive skills. The instrumental variables include the employment and wage in the top destination provinces provided in the China Labor Statistical Yearbook. The results are based on FE with IV specification used in column (4) and (8) in Table 3.7.

Table 3.19: Father's Migration Status and Child Non-cognitive Skills: Value-added Model

	Internalizing Score 2004		Externalizing Score 2004	
	OLS	2SLS	OLS	2SLS
Father Migration	-0.060**	-0.082***	-0.079**	-0.080***
	(0.025)	(0.035)	(0.028)	(0.034)
Internalizing Score 2000	0.040	0.040		
	(0.029)	(0.029)		
Externalizing Score 2000			0.128***	0.128***
			(0.030)	(0.029)
F test		15.78		15.78
Over-identification test		0.43		0.43
Observations	1779	1779	1779	1779
R^2	0.052	0.034	0.067	0.049

Note: Standard errors in brackets are significant at *10% **5%, and ***1% and clustered at the county level. The dependent variables are the internalizing and externalizing scores. Higher internalizing or externalizing scores indicate higher non-cognitive skills. The instrumental variables include the employment and wage in the top destination provinces provided in the China Labor Statistical Yearbook. For OLS and 2SLS specifications, control variables include child's age and gender, parental age and education, number of siblings, whether any grandparent resides at home or not, village level population, whether has elementary school, and whether has middle school, as well as county and year fixed effects.

Table 3.20: Mechanisms: Parenting Style and Parent Health Condition (FE+IV)

	Harsh Parenting		Health Condition	
	Mother	Child	Father	Mother
Father Migration	-0.388	0.071	0.079	-0.200
	(0.332)	(0.370)	(0.244)	(0.257)
F test	69.22	69.71	68.68	69.65
Observations	3504	3556	3504	3542
RMSE	0.535	0.607	0.395	0.421

Note: Parenting style is measured as whether the parents beat the children when children misbehave. Both the mother and the target child have answered these questions. The father's and mother's health conditions are self-reported. Standard errors in brackets are significant at *10% **5%, and ***1% and clustered at the county level. The results are based on FE with IV specification used in column (4) and (8) in Table 3.7.

Bibliography

Abel, Martin, Rulof Burger, Eliana Carranza, and Patrizio Piraino, "Bridging the intention-behavior gap? The effect of plan-making prompts on job search and employment," *American Economic Journal: Applied Economics*, 2019, *11* (2), 284–301.

Acosta, Pablo, "School Attendance, Child Labour, and Remittances from International Migration in El Salvador," *Journal of Development Studies*, 2011, *47* (6), 913–936.

Akee, Randall, William Copeland, E Jane Costello, and Emilia Simeonova, "How does household income affect child personality traits and behaviors?," *American Economic Review*, 2018, *108* (3), 775–827.

Alcaraz, Carlo, Daniel Chiquiar, and Alejandrina Salcedo, "Remittances, Schooling, and Child Labor in Mexico," *Journal of Development Economics*, 2012, *97* (1), 156–165.

BIBLIOGRAPHY

Almlund, Mathilde, Angela Lee Duckworth, James Heckman, and Tim Kautz, "Personality psychology and economics," in "Handbook of the Economics of Education," Vol. 4, Elsevier, 2011, pp. 1–181.

Altintas, Evrim, "The widening education gap in developmental child care activities in the United States, 1965–2013," *Journal of Marriage and Family*, 2016, *78* (1), 26–42.

Amuedo-Dorantes, Catalina and Susan Pozo, "Accounting for remittance and migration effects on children's schooling," *World Development*, 2010, *38* (12), 1747–1759.

Andrich, David, "A Rating Formulation for Ordered Response Categories," *Psychometrika*, 1978, *43* (4), 561–573.

Angrist, Joshua D, Guido W Imbens, and Donald B Rubin, "Identification of causal effects using instrumental variables," *Journal of the American statistical Association*, 1996, *91* (434), 444–455.

Antman, Francisca M, "Adult Child Migration and the Health of Elderly Parents Left Behind in Mexico," *American Economic Review*, 2010, *100* (2), 205–08.

_ , "The Intergenerational Effects of Paternal Migration on Schooling and Work: What Can We Learn from Children's Time Allocations?," *Journal of Development Economics*, 2011, *96* (2), 200–208.

BIBLIOGRAPHY

_ , "Gender, educational attainment, and the impact of parental migration on children left behind," *Journal of Population Economics*, 2012, *25* (4), 1187–1214.

_ , "The impact of migration on family left behind," in "International handbook on the economics of migration," Edward Elgar Publishing, 2013.

_ , "Spousal Employment and Intra-household Bargaining Power," *Applied Economics Letters*, 2014, *21* (8), 560–563.

Arni, Patrick and Amelie Schiprowski, "Job search requirements, effort provision and labor market outcomes," *Journal of Public Economics*, 2019, *169*, 65–88.

Arnott, Richard, "Time for revisionism on rent control?," *Journal of Economic Perspectives*, 1995, *9* (1), 99–120.

Asquith, Brian J., "Rent Control and Evictions: Evidence from San Francisco," *Job Market Paper*, 2017.

Atella, Vincenzo, Partha Deb, and Joanna Kopinska, "Heterogeneity in Long Term Health Outcomes of Migrants Within Italy," *Journal of Health Economics*, 2019, *63*, 19–33.

Ault, Richard W, John D Jackson, and Richard P Saba, "The effect of long-term rent control on tenant mobility," *Journal of Urban Economics*, 1994, *35* (2), 140–158.

BIBLIOGRAPHY

Austin, Benjamin, Edward Glaeser, and Lawrence Summers, "Jobs for the Heartland: Place-Based Policies in 21st-century America," *Brookings Papers on Economic Activity*, 2018, pp. 151–233.

Autor, David, Christopher J Palmer, and Parag A Pathak, "Gentrification and the Amenity Value of Crime Reductions: Evidence from Rent Deregulation," Technical Report, National Bureau of Economic Research 2017.

Autor, David H, Christopher J Palmer, and Parag A Pathak, "Housing market spillovers: Evidence from the end of rent control in Cambridge, Massachusetts," *Journal of Political Economy*, 2014, *122* (3), 661–717.

Averett, Susan and Yang Wang, "The effects of earned income tax credit payment expansion on maternal smoking," *Health Economics*, 2013, *22* (11), 1344–1359.

Bai, Yu, Linxiu Zhang, Chengfang Liu, Yaojiang Shi, Di Mo, and Scott Rozelle, "Effect of Parental Migration on the Academic Performance of Left behind Children in North Western China," *The Journal of Development Studies*, 2017, pp. 1–17.

Barrett, Alan and Irene Mosca, "The Psychic Costs of Migration: Evidence from Irish Return Migrants," *Journal of Population Economics*, 2013, *26* (2), 483–506.

BIBLIOGRAPHY

Basten, Christoph, Andreas Fagereng, and Kjetil Telle, "Cash-on-hand and the Duration of Job Search: Quasi-experimental Evidence from Norway," *The Economic Journal*, 2014, *124* (576), 540–568.

Bastian, Jacob and Katherine Michelmore, "The long-term impact of the earned income tax credit on children's education and employment outcomes," *Journal of Labor Economics*, 2018, *36* (4), 1127–1163.

Belot, Michele and Jan C. van Ours, "Unemployment and labor market institutions: An empirical analysis," *Journal of The Japanese and International Economies*, 2001, *15* (4), 403–418.

Bernal, Raquel, "The effect of maternal employment and child care on children's cognitive development," *International Economic Review*, 2008, *49* (4), 1173–1209.

_ **and Michael P Keane**, "Quasi-structural estimation of a model of childcare choices and child cognitive ability production," *Journal of Econometrics*, 2010, *156* (1), 164–189.

_ **and** _ , "Child care choices and children's cognitive achievement: The case of single mothers," *Journal of Labor Economics*, 2011, *29* (3), 459–512.

Bertrand, Marianne and Jessica Pan, "The Trouble with Boys: Social Influences and the Gender Gap in Disruptive Behavior," *American Economic Journal: Applied Economics*, 2013, *5* (1), 32–64.

BIBLIOGRAPHY

Bhuller, Manudeep, Gordon B Dahl, Katrine Vellesen Loken, and Magne Mogstad, "Incarceration, Recidivism, and Employment," Technical Report, IZA Discussion Papers 2018.

Binci, Michele and Gianna Claudia Giannelli, "Internal versus International Migration: Impacts of Remittances on Child Labor and Schooling in Vietnam," *International Migration Review*, 2018, *52* (1), 43–65.

Blanchflower, David G and Andrew J Oswald, "Does high home-ownership impair the labor market?," Technical Report, National Bureau of Economic Research 2013.

Boca, Daniela Del, Chiara Monfardini, and Cheti Nicoletti, "Parental and child time investments and the cognitive development of adolescents," *Journal of Labor Economics*, 2017, *35* (2), 565–608.

_ , Christopher Flinn, and Matthew Wiswall, "Household Choices and Child Development," *The Review of Economic Studies*, 2014, *81* (1), 137–185.

Bono, Emilia Del, Marco Francesconi, Yvonne Kelly, and Amanda Sacker, "Early Maternal Time Investment and Early Child Outcomes," *Economic Journal*, October 2016, *126*, F96–135.

BIBLIOGRAPHY

Bono, Emilia Del, Marco Francesconi, Yvonne Kelly, and Amanda Sacker, "Early maternal time investment and early child outcomes," *The Economic Journal*, 2016, *126* (596), F96–F135.

Borghans, Lex, Angela Lee Duckworth, James J Heckman, and Bas Ter Weel, "The economics and psychology of personality traits," *Journal of human Resources*, 2008, *43* (4), 972–1059.

Botezat, Alina and Friedhelm Pfeiffer, "The Impact of Parents Migration on the Well-Being of Children Left Behind–Initial Evidence from Romania," *ZEW-Centre for European Economic Research Discussion Paper*, 2014, (14-029).

Bound, John and Harry J Holzer, "Demand Shifts, Population Adjustments, and Labor Market Outcomes during the 1980s," *Journal of Labor Economics*, 2000, *18* (1), 20–54.

_ **, David A Jaeger, and Regina M Baker**, "Problems with Instrumental Variables Estimation When the Correlation Between the Instruments and the Endogeneous Explanatory Variable is Weak," *Journal of the American Statistical Association*, 1995, pp. 443–450.

Boyd-Swan, Casey, Chris M Herbst, John Ifcher, and Homa Zarghamee, "The earned income tax credit, mental health, and happiness," *Journal of Economic Behavior & Organization*, 2016, *126*, 18–38.

BIBLIOGRAPHY

Brauw, Alan De and Ren Mu, "Migration and the Overweight and Underweight Status of Children in Rural China," *Food Policy*, 2011, *36* (1), 88–100.

Caliendo, Marco, Steffen Künn, and Robert Mahlstedt, "The return to labor market mobility: An evaluation of relocation assistance for the unemployed," *Journal of Public Economics*, 2017, *148*, 136–151.

Card, David, "Estimating the return to schooling: Progress on some persistent econometric problems," *Econometrica*, 2001, *69* (5), 1127–1160.

_ , Raj Chetty, and Andrea Weber, "Cash-on-Hand and Competing Models of Intertemporal Behavior: New Evidence from the Labor Market," *The Quarterly Journal of Economics*, 2007, *122* (4), 1511–1560.

Carneiro, Pedro, Costas Meghir, and Matthias Parey, "Maternal education, home environments, and the development of children and adolescents," *Journal of the European Economic Association*, 2013, *11* (suppl_1), 123–160.

Charles, Kerwin Kofi, Erik Hurst, and Nikolai Roussanov, "Conspicuous consumption and race," *The Quarterly Journal of Economics*, 2009, *124* (2), 425–467.

Chen, Joyce J, "Identifying Non-cooperative Behavior Among Spouses: Child Outcomes in Migrant-sending Households," *Journal of Development Economics*, 2013, *100* (1), 1–18.

BIBLIOGRAPHY

Chernozhukov, Victor, Iván Fernández-Val, and Ye Luo, "The sorted effects method: discovering heterogeneous effects beyond their averages," *Econometrica*, 2018, *86* (6), 1911–1938.

Chesher, Andrew, "Identification in nonseparable models," *Econometrica*, 2003, *71* (5), 1405–1441.

Chetty, Raj, "Moral Hazard versus Liquidity and Optimal Unemployment Insurance," *Journal of Political Economy*, 2008, *116* (2), 173–234.

_ , John N Friedman, and Emmanuel Saez, "Using Differences in Knowledge across Neighborhoods to Uncover the Impacts of the EITC on Earnings," *American Economic Review*, 2013, *103* (7), 2683–2721.

_ , _ , and Jonah E Rockoff, "New evidence on the long-term impacts of tax credits," *IRS Statistics of Income White Paper*, 2011.

_ , Nathaniel Hendren, and Lawrence F Katz, "The effects of exposure to better neighborhoods on children: New evidence from the moving to opportunity experiment," *American Economic Review*, 2016, *106* (4), 855–902.

Chyn, Eric, "Moved to opportunity: The long-run effects of public housing demolition on children," *American Economic Review*, 2018, *108* (10), 3028–56.

Cingano, Federico and Alfonso Rosolia, "People I know: job search and social networks," *Journal of Labor Economics*, 2012, *30* (2), 291–332.

BIBLIOGRAPHY

Collinson, Robert, Ingrid Gould Ellen, and Jens Ludwig, "Low-Income Housing Policy," in "Economics of Means-Tested Transfer Programs in the United States, volume 2," University of Chicago Press, 2015, pp. 59–126.

Cortes, Patricia, "The Feminization of International Migration and its Effects on the Children Left Behind: Evidence from the Philippines," *World Development*, 2015, *65*, 62–78.

Coulson, N Edward and Lynn M Fisher, "Tenure choice and labour market outcomes," *Housing Studies*, 2002, *17* (1), 35–49.

_ **and** _ , "Housing tenure and labor market impacts: The search goes on," *Journal of Urban Economics*, 2009, *65* (3), 252–264.

Cowan, Benjamin and Nathan Tefft, "Education, Maternal Smoking, and the Earned Income Tax Credit," *The BE Journal of Economic Analysis & Policy*, 2012, *12* (1), 1–39.

Cox-Edwards, Alejandra and Manuelita Ureta, "International Migration, Remittances, and Schooling: Evidence from El Salvador," *Journal of Development Economics*, 2003, *72* (2), 429–461.

Cunha, Flavio and James J Heckman, "Formulating, identifying and estimating the technology of cognitive and noncognitive skill formation," *Journal of Human Resources*, 2008, *43* (4), 738–782.

BIBLIOGRAPHY

_ , _ , **and Susanne M Schennach**, "Estimating the technology of cognitive and noncognitive skill formation," *Econometrica*, 2010, *78* (3), 883–931.

Currie, Janet and Douglas Almond, "Human capital development before age five," *Handbook of Labor Economics*, 2011, *4*, 1315–1486.

Dahl, Gordon B and Lance Lochner, "The impact of family income on child achievement: Evidence from the earned income tax credit," *The American Economic Review*, 2012, *102* (5), 1927–1956.

_ **and** _ , "The Impact of Family Income on Child Achievement: Evidence from the Earned Income Tax Credit: Reply," *The American Economic Review*, 2017, *107* (2), 629–631.

de Brauw, Alan and John Giles, "Migrant Opportunity and the Educational Attainment of Youth in Rural China," *Journal of Human Resources*, 2017, *52* (1), 272–311.

DellaVigna, Stefano and M Daniele Paserman, "Job search and impatience," *Journal of Labor Economics*, 2005, *23* (3), 527–588.

Diamond, Rebecca, Tim McQuade, and Franklin Qian, "The effects of rent control expansion on tenants, landlords, and inequality: Evidence from San Francisco," *American Economic Review*, 2019, *109* (9), 3365–94.

BIBLIOGRAPHY

Dickson, Matt, Paul Gregg, and Harriet Robinson, "Early, Late or Never? When does Parental Education Impact Child Outcomes?," *The Economic Journal*, 2016, *126* (596), F184–F231.

Dijk, Winnie Van, "The socio-economic consequences of housing assistance," Technical Report, Working Paper 2019.

Doepke, Matthias and Fabrizio Zilibotti, "Parenting with Style: Altruism and Paternalism in Intergenerational Preference Transmission," *Econometrica*, 2017, *85* (5), 1331–1371.

Dooley, Martin and Jennifer Stewart, "Family income, parenting styles and child behavioural–emotional outcomes," *Health economics*, 2007, *16* (2), 145–162.

Du, Yang, Albert Park, and Sangui Wang, "Migration and Rural Poverty in China," *Journal of Comparative Economics*, 2005, *33* (4), 688–709.

Duffy, Denise and Narayan Sastry, "Achievement Tests in the Panel Study of Income Dynamics Child Development Supplement," *PSID Technical Series Papers #14-02*, 2014.

Early, Dirk W, "Rent control, rental housing supply, and the distribution of tenant benefits," *Journal of Urban Economics*, 2000, *48* (2), 185–204.

_ **and Edgar O Olsen**, "Rent control and homelessness," *Regional Science and Urban Economics*, 1998, *28* (6), 797–816.

BIBLIOGRAPHY

Edin, Kathryn and Laura Lein, *Making ends meet: How single mothers survive welfare and low-wage work*, Russell Sage Foundation, 1997.

Eissa, Nada and Jeffrey B Liebman, "Labor Supply Response to the Earned Income Tax Credit," *The Quarterly Journal of Economics*, 1996, *111* (2), 605–637.

Ermisch, John and Marco Francesconi, "The effect of parental employment on child schooling," *Journal of Applied Econometrics*, 2013, *28* (5), 796–822.

Esquivel, Gerardo and Alejandra Huerta-Pineda, "Remittances and Poverty in Mexico: a Propensity Score Matching Approach," *Integration and Trade Journal*, 2007, *27* (11), 45–57.

Evans, William N and Craig L Garthwaite, "Giving mom a break: The impact of higher EITC payments on maternal health," *American Economic Journal: Economic Policy*, 2014, *6* (2), 258–290.

Fan, Jessie X and Joan Koonce Lewis, "Budget allocation patterns of African Americans," *Journal of Consumer Affairs*, 1999, *33* (1), 134–164.

Favilukis, Jack Y, Pierre Mabille, and Stijn Van Nieuwerburgh, "Affordable Housing and City Welfare," *Columbia Business School Research Paper*, 2019, (18-77).

BIBLIOGRAPHY

Feller, Avi, Todd Grindal, Luke Miratrix, Lindsay C Page et al., "Compared to what? Variation in the impacts of early childhood education by alternative care type," *The Annals of Applied Statistics*, 2016, *10* (3), 1245–1285.

Figlio, David, Krzysztof Karbownik, Jeffrey Roth, Melanie Wasserman et al., "Family Disadvantage and the Gender Gap in Behavioral and Educational Outcomes," *American Economic Journal: Applied Economics*, 2019, *11* (3), 338–81.

Fiorini, Mario and Michael P Keane, "How the allocation of children's time affects cognitive and noncognitive development," *Journal of Labor Economics*, 2014, *32* (4), 787–836.

Flossmann, A, R Piatek, and L Wichert, "Learning for Life from Schooling: The Role of Educational Institutions in Non-Cognitive Skill Formation," *unpublished working paper*, 2006.

Folbre, Nancy, Jayoung Yoon, Kade Finnoff, and Allison Sidle Fuligni, "By what measure? Family time devoted to children in the United States," *Demography*, 2005, *42* (2), 373–390.

Ganong, Peter and Daniel Shoag, "Why has regional income convergence in the US declined?," *Journal of Urban Economics*, 2017, *102*, 76–90.

BIBLIOGRAPHY

Gao, Yang, Li Ping Li, Jean Hee Kim, Nathan Congdon, Joseph Lau, and Sian Griffiths, "The Impact of Parental Migration on Health Status and Health Behaviours among Left Behind Adolescent School Children in China," *BMC public Health*, 2010, *10* (1), 56.

Giannelli, Gianna Claudia and Lucia Mangiavacchi, "Children's Schooling and Parental Migration: Empirical Evidence on the 'Left-behind' Generation in Albania," *Labour*, 2010, *24*, 76–92.

Gibson, John and David McKenzie, "The Development Impact of a Best Practice Seasonal Worker Policy," *Review of Economics and Statistics*, 2014, *96* (2), 229–243.

_ , _ , and Steven Stillman, "The Impacts of International Migration on Remaining Household Members: Omnibus Results from a Migration Lottery Program," *Review of Economics and Statistics*, 2011, *93* (4), 1297–1318.

_ , _ , and _ , "What Happens to Diet and Child Health When Migration Splits Households? Evidence from a Migration Lottery Program," *Food Policy*, 2011, *36* (1), 7–15.

Ginther, Donna K and Robert A Pollak, "Family Structure and Children's Educational Outcomes: Blended Families, Stylized Facts, and Descriptive Regressions," *Demography*, 2004, *41* (4), 671–696.

BIBLIOGRAPHY

Glaeser, Edward L, "The social costs of rent control revisited," Technical Report, National Bureau of Economic Research 1996.

_ **and Erzo FP Luttmer**, "The Misallocation of Housing under Rent Control," *American Economic Review*, 2003, *93* (4), 1027–1046.

Glewwe, Paul, Albert Park, and Meng Zhao, "A better Vision for Development: Eyeglasses and Academic Performance in Rural Primary Schools in China," *Journal of Development Economics*, 2016, *122*, 170–182.

_ **, Qiuqiong Huang, and Albert Park**, "Cognitive Skills, Noncognitive Skills, and School-to-work Transitions in Rural China," *Journal of Economic Behavior & Organization*, 2017, *134*, 141–164.

Goodman, David SG, "The Campaign to "Open up the West": National, Provincial-level and Local Perspectives," *The China Quarterly*, 2004, *178*, 317–334.

Green, Richard K and Patric H Hendershott, "Home-ownership and unemployment in the US," *Urban Studies*, 2001, *38* (9), 1509–1520.

_ **and Stephen Malpezzi**, *A primer on US housing markets and housing policy*, The Urban Insitute, 2003.

BIBLIOGRAPHY

Grogger, Jeffrey, "The Effects of Time Limits, the EITC, and Other Policy Changes on Welfare Use, Work, and Income among Female-Headed Families," *Review of Economics and Statistics*, 2003, pp. 394–408.

Gyourko, Joseph and Edward Glaeser, "Rethinking Federal Housing Policy," Technical Report, American Enterprise Institute 2008.

_ **and Peter Linneman**, "Equity and efficiency aspects of rent control: An empirical study of New York City," *Journal of Urban Economics*, 1989, *26* (1), 54–74.

_ **and** _ , "Rent controls and rental housing quality: A note on the effects of New York City's old controls," *Journal of Urban Economics*, 1990, *27* (3), 398–409.

_ **, Christopher Mayer, and Todd Sinai**, "Superstar cities," *American Economic Journal: Economic Policy*, 2013, *5* (4), 167–99.

Hannum, Emily and Yuping Zhang, "Poverty and Proximate Barriers to Learning: Vision Deficiencies, Vision Correction and Educational Outcomes in Rural Northwest China," *World Development*, 2012, *40* (9), 1921–1931.

Hanson, Gordon H and Christopher Woodruff, "Emigration and Educational Attainment in Mexico," *Working Paper*, 2003.

Hao, Lingxin and Wei-Jun Jean Yeung, "Parental spending on school-age children: Structural stratification and parental expectation," *Demography*, 2015, *52* (3), 835–860.

BIBLIOGRAPHY

Heckman, James, "Instrumental variables: A study of implicit behavioral assumptions used in making program evaluations," *Journal of Human Resources*, 1997, pp. 441–462.

Heckman, James J and Edward Vytlacil, "Policy-relevant treatment effects," *American Economic Review*, 2001, *91* (2), 107–111.

_ **and** _ , "Structural equations, treatment effects, and econometric policy evaluation 1," *Econometrica*, 2005, *73* (3), 669–738.

_ **and Stefano Mosso**, "The Economics of Human Development and Social Mobility," *Annual Review of Economics*, 2014, *6* (1), 689–733.

_ **and Yona Rubinstein**, "The Importance of Noncognitive Skills: Lessons from the GED Testing Program," *American Economic Review*, 2001, *91* (2), 145–149.

_ **, Jora Stixrud, and Sergio Urzua**, "The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior," *Journal of Labor Economics*, 2006, *24* (3), 411–482.

Heckman, James, Rodrigo Pinto, and Peter Savelyev, "Understanding the mechanisms through which an influential early childhood program boosted adult outcomes," *The American Economic Review*, 2013, *103* (6), 2052–2086.

BIBLIOGRAPHY

Herbst, Chris M, "The impact of non-parental child care on child development: Evidence from the summer participation "dip"," *Journal of Public Economics*, 2013, *105*, 86–105.

_ , "Are Parental Welfare Work Requirements Good for Disadvantaged Children? Evidence From Age-of-Youngest-Child Exemptions," *Journal of Policy Analysis and Management*, 2017, *36* (2), 327–357.

_ **and Erdal Tekin**, "Child care subsidies and child development," *Economics of Education Review*, 2010, *29* (4), 618–638.

_ **and** _ , "The impact of child-care subsidies on child development: Evidence from geographic variation in the distance to social service agencies," *Journal of Policy Analysis and Management*, 2016, *35* (1), 94–116.

Hildebrandt, Nicole and David J McKenzie, "The Effects of Migration on Child Health in Mexico," *Economía*, 2005, *6* (1), 257–289.

Hoddinott, John and Lawrence Haddad, "Does Female Income Share Influence Household Expenditures? Evidence from Côte d'Ivoire," *oxford Bulletin of Economics and Statistics*, 1995, *57* (1), 77–96.

Hotz, V Joseph and John Karl Scholz, "The Earned Income Tax Credit," *Means-Tested Transfer Programs in the United States*, 2003, pp. 141–198.

BIBLIOGRAPHY

Hotz, V. Joseph and Matthew Wiswall, "Child Care and Child Care Policy: Existing Policies, Their Effects, and Reforms," *working paper*, 2019.

Hoynes, Hilary, Doug Miller, and David Simon, "Income, the earned income tax credit, and Infant health," *American Economic Journal: Economic Policy*, 2015, *7* (1), 172–211.

Hsin, Amy and Christina Felfe, "When does time matter? Maternal employment, children's time with parents, and child development," *Demography*, 2014, *51* (5), 1867–1894.

Hu, Feng, "Migration, Remittances, and Children's High School Attendance: The Case of Rural China," *International Journal of Educational Development*, 2012, *32* (3), 401–411.

Humphries, John Eric and Fabian Kosse, "On the Interpretation of Non-cognitive Skills–What is Being measured and Why it Matters," *Journal of Economic Behavior & Organization*, 2017, *136*, 174–185.

Imbens, Guido W, "Instrumental Variables: An Econometrician's Perspective," *Statistical Science*, 2014, pp. 323–358.

_ **and Joshua D Angrist**, "Identification and Estimation of Local Average Treatment Effects," *Econometrica*, 1994, *62* (2), 467–475.

BIBLIOGRAPHY

_ **and Whitney K Newey**, "Identification and estimation of triangular simultaneous equations models without additivity," *Econometrica*, 2009, *77* (5), 1481–1512.

Jacob, Brian A and Jens Ludwig, "The effects of housing assistance on labor supply: Evidence from a voucher lottery," *American Economic Review*, 2012, *102* (1), 272–304.

Jiang, Hanchen, Luis E. Quintero, and Xi Yang, "Housing Policy, Centrality and Mobility: The case of Manhattan in New York City," 2019.

Kalil, Ariel, Rebecca Ryan, and Michael Corey, "Diverging destinies: Maternal education and the developmental gradient in time with children," *Demography*, 2012, *49* (4), 1361–1383.

Kautz, Tim, James J Heckman, Ron Diris, Bas Ter Weel, and Lex Borghans, "Fostering and Measuring Skills: Improving Cognitive and Non-cognitive Skills to Promote Lifetime Success," *NBER Working Paper*, 2014.

Kearney, Melissa S. and Magne Mogstad, "Universal Basic Income (UBI) as a Policy Response to Current Challenges," Technical Report, Economic Strategy Group at The Aspen Institute 2019.

Kessler, Ronald C, Peggy R Barker, Lisa J Colpe, Joan F Epstein, Joseph C Gfroerer, Eva Hiripi, Mary J Howes, Sharon-Lise T Normand, Ronald W Manderscheid, Ellen E Walters et al., "Screening for Serious

Mental Illness in the General Population," *Archives of General Psychiatry*, 2003, *60* (2), 184–189.

Kline, Patrick and Christopher R Walters, "Evaluating public programs with close substitutes: The case of Head Start," *The Quarterly Journal of Economics*, 2016, *131* (4), 1795–1848.

Kooreman, Peter, "Time, Money, Peers, and Parents; Some Data and Theories on Teenage Behavior," *Journal of Population Economics*, 2007, *20* (1), 9–33.

Kroeger, Antje and Kathryn H Anderson, "Remittances and the Human Capital of Children: New Evidence from Kyrgyzstan during Revolution and Financial Crisis, 2005–2009," *Journal of Comparative Economics*, 2014, *42* (3), 770–785.

Lee, Leng and Albert Park, "Parental Migration and Child Development in China," 2010.

Leight, Jessica, "Sibling rivalry: Endowment and intrahousehold allocation in Gansu Province, China," *Economic Development and Cultural Change*, 2017, *65* (3), 457–493.

_ **and Elaine M Liu**, "Maternal Education, Parental Investment and Non-Cognitive Skills in Rural China," Technical Report, National Bureau of Economic Research 2016.

BIBLIOGRAPHY

Lépine, Aurélia and Eric Strobl, "The Effect of Women's Bargaining Power on Child Nutrition in Rural Senegal," *World Development*, 2013, *45*, 17–30.

Li, Lixing and Xiaoyu Wu, "Gender of Children, Bargaining Power, and Intrahousehold Resource Allocation in China," *Journal of Human Resources*, 2011, *46* (2), 295–316.

Lindqvist, Erik and Roine Vestman, "The Labor Market Returns to Cognitive and Noncognitive Ability: Evidence from the Swedish Enlistment," *American Economic Journal: Applied Economics*, 2011, *3* (1), 101–28.

Linneman, Peter, "The effect of rent control on the distribution of income among New York City renters," *Journal of Urban Economics*, 1987, *22* (1), 14–34.

Lippman, Steven A and John J McCall, "The Economics of Job Search: A Survey," *Economic Inquiry*, 1976, *14* (2), 155.

Liu, Haiyong, Thomas A Mroz, and Wilbert Van der Klaauw, "Maternal employment, migration, and child development," *Journal of Econometrics*, 2010, *156* (1), 212–228.

Liu, Zhengkui, Xinying Li, and Xiaojia Ge, "Left too Early: the Effects of Age at Separation from Parents on Chinese Rural Children's Symptoms of Anxiety and Depression," *American Journal of Public Health*, 2009, *99* (11), 2049–2054.

BIBLIOGRAPHY

Lundberg, Shelly, Jennifer L Romich, and Kwok Ping Tsang, "Decision-making by children," *Review of Economics of the Household*, 2009, *7* (1), 1–30.

Lundstrom, Samuel, "The impact of family income on child achievement: Evidence from the Earned Income Tax Credit: Comment," *The American Economic Review*, 2017, *107* (2), 623–628.

Macours, Karen and Renos Vakis, "Seasonal Migration and Early Childhood Development," *World Development*, 2010, *6* (38), 857–869.

Manoli, Day and Nicholas Turner, "Cash-on-Hand and College Enrollment: Evidence from Population Tax Data and the Earned Income Tax Credit," *American Economic Journal: Economic Policy*, 2018, *10* (2), 242–71.

Mansuri, Ghazala, "Migration, School Attainment, and Child Labor: Evidence from Rural Pakistan," *Working Paper*, 2006.

Maxfield, Michelle, "The effects of the Earned Income Tax Credit on child achievement and long-term educational attainment," *Michigan State University Job Market Paper*, 2013.

McCall, John Joseph, "Economics of information and job search," *The Quarterly Journal of Economics*, 1970, pp. 113–126.

BIBLIOGRAPHY

McKenzie, David and Hillel Rapoport, "Can Migration Reduce Educational Attainment? Evidence from Mexico," *Journal of Population Economics*, 2011, *24* (4), 1331–1358.

McLanahan, Sara, Laura Tach, and Daniel Schneider, "The Causal Effects of Father Absence," *Annual Review of Sociology*, 2013, *39*, 399–427.

Meng, Xin and Chikako Yamauchi, "Children of Migrants: The Cumulative Impact of Parental Migration on Children's Education and Health Outcomes in China," *Demography*, 2017, *54* (5), 1677–1714.

Mergo, Teferi, "The Effects of International Migration on Migrant-source Households: Evidence from Ethiopian Diversity-visa Lottery Migrants," *World Development*, 2016, *84*, 69–81.

Metcalf, Gabriel, "Sand castles before the tide? Affordable housing in expensive cities," *Journal of Economic Perspectives*, 2018, *32* (1), 59–80.

Meyer, Bruce D and James X Sullivan, "Changes in the consumption, income, and well-being of single mother headed families," *American Economic Review*, 2008, *98* (5), 2221–41.

Meyerhoefer, Chad D and CJ Chen, "The Effect of Parental Labor Migration on Children's Educational Progress in Rural China," *Review of Economics of the Household*, 2011, *9* (3), 379–396.

BIBLIOGRAPHY

Michelmore, Katherine, "The effect of income on educational attainment: Evidence from state earned income tax credit expansions," 2013.

Moffitt, Robert A, "Unemployment benefits and unemployment," *IZA World of Labor*, 2014, pp. 1–13.

Mogstad, Magne and Alexander Torgovitsky, "Identification and extrapolation of causal effects with instrumental variables," *Annual Review of Economics*, 2018, *10*, 577–613.

_ , Andres Santos, and Alexander Torgovitsky, "Using instrumental variables for inference about policy relevant treatment parameters," *Econometrica*, 2018, *86* (5), 1589–1619.

Moon, Choon-Geol and Janet G Stotsky, "The effect of rent control on housing quality change: a longitudinal analysis," *Journal of Political Economy*, 1993, *101* (6), 1114–1148.

Morescalchi, Andrea, "The puzzle of job search and housing tenure: A reconciliation of theory and empirical evidence," *Journal of Regional Science*, 2016, *56* (2), 288–312.

Morris, Pamela A, Maia Connors, Allison Friedman-Krauss, Dana Charles McCoy, Christina Weiland, Avi Feller, Lindsay Page, Howard Bloom, and Hirokazu Yoshikawa, "New findings on impact variation from the Head

Start Impact Study: Informing the scale-up of early childhood programs," *AERA Open*, 2018, *4* (2), 2332858418769287.

Mu, Ren and Alan De Brauw, "Migration and Young Child Nutrition: Evidence from Rural China," *Journal of Population Economics*, 2015, *28* (3), 631–657.

Munch, Jakob Roland, Michael Rosholm, and Michael Svarer, "Are homeowners really more unemployed?," *The Economic Journal*, 2006, *116* (514), 991–1013.

_ , _ , and _ , "Home ownership, job duration, and wages," *Journal of Urban Economics*, 2008, *63* (1), 130–145.

Murray, Michael P, "Avoiding Invalid Instruments and Coping with Weak Instruments," *Journal of Economic Perspectives*, 2006, *20* (4), 111–132.

Nagy, John, "Increased Duration and Sample Attrition in New York City's Rent Controlled Sector," *Journal of Urban Economics*, 1995, *38* (2), 127–137.

_ , "Do Vacancy Decontrol Provisions Undo Rent Control?," *Journal of Urban Economics*, 1997, *42* (1), 64–78.

Newman, Kathe and Elvin K Wyly, "The Right to Stay Put, Revisited: Gentrification and Resistance to Displacement in New York City," *Urban Studies*, 2006, *43* (1), 23–57.

BIBLIOGRAPHY

Newman, Sandra, C Scott Holupka, and Joseph Harkness, "The long-term effects of housing assistance on work and welfare," *Journal of Policy Analysis and Management: The Journal of the Association for Public Policy Analysis and Management*, 2009, *28* (1), 81–101.

Nichols, Austin and Jesse Rothstein, "The earned income tax credit," in "Economics of Means-Tested Transfer Programs in the United States, Volume 1," University of Chicago Press, 2015, pp. 137–218.

Nobles, Jenna, "Parenting from Abroad: Migration, Nonresident Father Involvement, and Children's Education in Mexico," *Journal of Marriage and Family*, 2011, *73* (4), 729–746.

Notowidigdo, Matthew J, "The Incidence of Local Labor Demand Shocks," 2019.

Olsen, Edgar O, "An econometric analysis of rent control," *Journal of Political Economy*, 1972, *80* (6), 1081–1100.

_ , "What do economists know about the effect of rent control on housing maintenance?," *The Journal of Real Estate Finance and Economics*, 1988, *1* (3), 295–307.

_ , "Housing programs for low-income households," in "Means-tested transfer programs in the United States," University of Chicago Press, 2003, pp. 365–442.

Oster, Emily, "Unobservable selection and coefficient stability: Theory and evidence," *Journal of Business & Economic Statistics*, 2019, *37* (2), 187–204.

BIBLIOGRAPHY

Oswald, Andrew, "The housing market and Europe's unemployment: a non-technical paper," *Homeownership and the labour Market in Europe*, 1999.

Oswald, Andrew J, "A conjecture on the explanation for high unemployment in the industrialized nations: Part I," Technical Report 1996.

Paserman, M Daniele, "Job search and hyperbolic discounting: Structural estimation and policy evaluation," *The Economic Journal*, 2008, *118* (531), 1418–1452.

Pastor, Manuel, Vanessa Carter, and Maya Abood, "Rent Matters: What are the Impacts of Rent Stabilization Measures?," 2018.

Powers, Elizabeth T, "The Impact of Economic Migration on Children's Cognitive Development: Evidence from the Mexican Family Life Survey," 2011.

Rapoport, Hillel and Frédéric Docquier, "The Economics of Migrants' Remittances," *Handbook of the Economics of Giving, Altruism and Reciprocity*, 2006, *2*, 1135–1198.

Ren, Qiang and Donald J Treiman, "The Consequences of Parental Labor Migration in China for Children's Emotional Wellbeing," *Social Science Research*, 2016, *58*, 46–67.

Ronda, Victor A. C., "The Effect of Maternal Psychological Distress on Children's Cognitive Development," *Job Market Paper*, 2016.

BIBLIOGRAPHY

Rosen, Sherwin, "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition," *The Journal of Political Economy*, 1974, *82* (1), 34–55.

Sandberg, John F and Sandra L Hofferth, "Changes in children's time with parents: United States, 1981–1997," *Demography*, 2001, *38* (3), 423–436.

Sargan, John D, "Testing for Misspecification after Estimating using Instrumental Variables," *Contributions to Econometrics: John Denis Sargan*, 1988, *1*, 213–235.

Sayer, LC, SM Bianchi, and JP Robinson, "Are Parents Investing Less in Children? Trends in Mothers' and Fathers' Time with Children.," *American Journal of Sociology*, 2004, *110* (1), 1–43.

Shroder, Mark, "Does housing assistance perversely affect self-sufficiency? A review essay," *Journal of Housing Economics*, 2002, *11* (4), 381–417.

Sieg, Holger and Chamna Yoon, "Waiting for Affordable Housing," 2016.

_ , V Kerry Smith, H Spencer Banzhaf, Randy Walsh et al., "Interjurisdictional housing prices in locational equilibrium," *Journal of Urban Economics*, 2002, *52* (1), 131–153.

Sims, David P, "Out of control: What can we learn from the end of Massachusetts rent control?," *Journal of Urban Economics*, 2007, *61* (1), 129–151.

Song, Yang, "What Should Economists Know about the Current Chinese hukou System?," *China Economic Review*, 2014, *29*, 200–212.

BIBLIOGRAPHY

Staiger, Douglas and James H Stock, "Instrumental Variables Regression with Weak Instruments," *Econometrica*, 1997, *65* (3), 557–586.

Stillman, Steven, John Gibson, and David McKenzie, "The Impact of Immigration on Child Health: Experimental Evidence from a Migration Lottery Program," *Economic Inquiry*, 2012, *50* (1), 62–81.

_ , _ , _ , and Halahingano Rohorua, "Miserable Migrants? Natural Experiment Evidence on International Migration and Objective and Subjective Well-being," *World Development*, 2015, *65*, 79–93.

Stock, James H, Jonathan H Wright, and Motohiro Yogo, "A survey of weak instruments and weak identification in generalized method of moments," *Journal of Business & Economic Statistics*, 2002, *20* (4), 518–529.

_ , Motohiro Yogo et al., "Testing for Weak Instruments in Linear IV Regression," Technical Report, National Bureau of Economic Research, Inc 2002.

Strully, Kate W, David H Rehkopf, and Ziming Xuan, "Effects of prenatal poverty on infant health: state earned income tax credits and birth weight," *American Sociological Review*, 2010, *75* (4), 534–562.

Su, Yaqin, Petros Tesfazion, and Zhong Zhao, "Where are the Migrants from? Inter-vs. Intra-provincial Rural-urban Migration in China," *China Economic Review*, 2018, *47*, 142–155.

BIBLIOGRAPHY

Susin, Scott, "Longitudinal outcomes of subsidized housing recipients in matched survey and administrative data," *Cityscape*, 2005, pp. 189–218.

Svarer, Michael, Michael Rosholm, and Jakob Roland Munch, "Rent control and unemployment duration," *Journal of Public Economics*, 2005, *89* (11), 2165–2181.

Tella, Rafael Di and Robert MacCulloch, "The consequences of labor market flexibility: Panel evidence based on survey data," *European Economic Review*, 2005, *49* (5), 1225–1259.

Terza, Joseph V, Anirban Basu, and Paul J Rathouz, "Two-stage residual inclusion estimation: Addressing endogeneity in health econometric modeling," *Journal of Health Economics*, 2008, *27* (3), 531–543.

Thiel, Hendrik and Stephan L Thomsen, "Noncognitive Skills in Economics: Models, Measurement, and Empirical Evidence," *Research in Economics*, 2013, *67* (2), 189–214.

Todd, Petra E and Kenneth I Wolpin, "On the specification and estimation of the production function for cognitive achievement," *The Economic Journal*, 2003, *113* (485), F3–F33.

_ **and** _ , "The production of cognitive achievement in children: Home, school, and racial test score gaps," *Journal of Human Capital*, 2007, *1* (1), 91–136.

BIBLIOGRAPHY

Topel, Robert H, "Local Labor Markets," *Journal of Political Economy*, 1986, *94* (3), 111–143.

_ **and Michael P Ward**, "Job mobility and the careers of young men," *The Quarterly Journal of Economics*, 1992, *107* (2), 439–479.

Turner, Bengt and Stephen Malpezzi, "A review of empirical evidence on the costs and benefits of rent control," *Swedish Economic Policy Review*, 2003, *10* (1), 11–60.

United Nations, "Trends in International Migrant Stock: The 2013 Revision," 2013.

Wang, Sophie Xuefei, "The Effect of Parental Migration on the Educational Attainment of Their Left-behind Children in Rural China," *The BE Journal of Economic Analysis & Policy*, 2014, *14* (3), 1037–1080.

Weinberger, Catherine J, "The Increasing Complementarity Between Cognitive and Social SKills," *Review of Economics and Statistics*, 2014, *96* (4), 849–861.

Wooldridge, Jeffrey M, "Control function methods in applied econometrics," *Journal of Human Resources*, 2015, *50* (2), 420–445.

World Bank, "From Poor Areas to Poor People: China's Evolving Poverty Reduction Agenda An Assessment of Poverty and Inequality in China," 2009.

BIBLIOGRAPHY

Xu, Hongwei and Yu Xie, "The causal effects of rural-to-urban migration on children's well-being in China," *European Sociological Review*, 2015, *31* (4), 502–519.

Yang, Dean, "International Migration, Remittances and Household Investment: Evidence from Philippine Migrants' Exchange Rate Shocks," *The Economic Journal*, 2008, *118* (528), 591–630.

_ , "Migrant Remittances," *Journal of Economic perspectives*, 2011, *25* (3), 129–52.

Yang, Xi, "The effects of home ownership on post-unemployment wages," *Regional Science and Urban Economics*, 2019, *74*, 1–17.

Zhang, Hongliang, Jere R Behrman, C Simon Fan, Xiangdong Wei, and Junsen Zhang, "Does Parental Absence Reduce Cognitive Achievements? Evidence from Rural China," *Journal of Development Economics*, 2014, *111*, 181–195.

Zhao, Qiran, Xiaohua Yu, Xiaobing Wang, and Thomas Glauben, "The Impact of Parental Migration on Children's School Performance in Rural China," *China Economic Review*, 2014, *31*, 43–54.

Zhao, Yaohui, "Leaving the Countryside: Rural-to-urban Migration Decisions in China," *American Economic Review*, 1999, pp. 281–286.

Curriculum Vitae

Hanchen Jiang graduated from Nanjing University with Bachelor in Economics in 2012. In 2014 he received a Master of Arts in Economics from the Graduate Program in Economic Development (GPED) at the Vanderbilt University. He enrolled in the Ph.D. program in economics at Johns Hopkins University in 2014 and will begin work as a tenure-track Assistant Professor in Economics at the University of North Texas in 2020 Fall.

His research is focused on the economics of low-income families and related economic policy issues in both developed countries and developing countries.